B3

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12Q 1/68, A61K 38/19, C07H 21/02, 21/04

A1

(11) International Publication Number: **WO 99/23253**

(43) International Publication Date: 14 May 1999 (14.05.99)

(21) International Application Number: PCT/US98/22578

(22) International Filing Date: 23 October 1998 (23.10.98)

(30) Priority Data:
60/063,832 30 October 1997 (30.10.97) US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application
US 60/063,832 (CON)
Filed on 30 October 1997 (30.10.97)

(71) Applicant *(for all designated States except US)*: THE GOVERNMENT OF THE UNITED STATES OF AMERICA, as represented by THE SECRETARY OF HEALTH AND HUMAN SERVICES [US/US]; National Institutes of Health, Office of Technology Transfer, Suite 325, 6011 Executive Boulevard, Rockville, MD 20852-3804 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: WINKLER, Cheryl, Ann [US/US]; 733 Tiffany Drive, Gaithersburg, MD 20878 (US). O'BRIEN, Stephen, J. [US/US]; 4201 Baker Valley Road, Frederick, MD 21702 (US).

(74) Agent: HAILE, Lisa, A.; Fish & Richardson P.C., Suite 1400, 4225 Executive Square, La Jolla, CA 92037 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: STROMAL CELL DERIVED FACTOR-1 (SDF-1) AND METHOD OF USE FOR DIAGNOSTIC AND PROGNOSTIC INDICATOR OF AIDS PATHOGENESIS

(57) Abstract

A nucleic acid sequence having a single nucleotide mutation in the 3' untranslated region of the mRNA transcript of the structural gene for stromal derived factor (*SDF1-3'A*) is provided. The mutation occured at an allele frequency of 6-26 % in various racial groups. SDF-1 is the principal ligand for CXCR4, a 7-transmembrane G-coupled receptor which, with CD4, provides an entry port for T-tropic HIV-1, a variety that frequently develops in AIDS patients just prior to CD4 T-lymphocyte depletion. Also provided in the invention is a method for determining the prognosis of a subject exposed to HIV-1 and a method for determining the susceptibility of a subject to HIV-1 infection.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

STROMAL CELL DERIVED FACTOR-1 (SDF-1) AND METHOD OF USE FOR DIAGNOSTIC AND PROGNOSTIC INDICATOR OF AIDS PATHOGENESIS

Field of the Invention

The present invention relates generally to HIV infection and more specifically to a stromal cell derived factor (SDF-1) variant that is associated with resistance to or decreased susceptibility to HIV infection.

Background of the Invention

The AIDS epidemic is characterized by considerable variation in the incidence of HIV-1 infection, in the rate of progression to CD4-T lymphocyte depletion, and in the development of AIDS . Part of the explanation for epidemiologic heterogeneity involves genetic variation in human loci that encode cellular factors which participate in HIV-1 infection and pathogenesis. For example, mutations in the *CCR5* and *CCR2* structural genes have been shown to be associated with delay in the median time required to develop AIDS, based on screens of mutant alleles in HIV-1 exposed cohorts (Samson, *Nature* 382:722, 1996; Dean *et al.*, *Science* 273:1856, 1996; Huang *et al.*, *Nature Med.* 2:1240, 1996; Michael *et al.*, *Nature Med.* 3:338, 1997; Zimmerman *et al.*, *Mol. Med.* 3:23, 1997; Biti *et al.*, *Nature Med.* 3:252, 1997; O'Brien *et al.*, *Lancet* 349:1219, 1997; Theodorou *et al.*, *Lancet* 349:1219, 1997; Smith *et al.*, *Science*, 277:959, 1997; D'Souza and V. A. Harden, *Nature Med.* 2:1293, 1996; B.A. Premack and Schall, *Nature Med.*, 2:1174 (1996); McNicholl *et al.*, *Emerg. Infect. Dis.* 3:261, 1997). The *CCR5-Δ32* deletion mutation specifies a reading frame shift that truncates the CCR5 protein, an obligate co-receptor with CD4 for HIV-1 infection of macrophages and monocytes which are the principal cell reservoir for early HIV-1 infection (Liu *et al.*, *Cell*

86:367, 1996; McNicholl *et al., Emerg. Infect. Dis.* 3:261, 1997; Alkhatib *et al., Science* 272:1955, 1996; Deng, *et al., Nature* 381:661, 1996; Dragic *et al., Nature* 381:667, 1996; Choe *et al., Cell* 85:1135, 1996; Doranz *et al., Cell* 85:1149, 1996; J. Rucker *et al., Cell* 87:437, 1996). *CCR5* deletion mutation homozygotes (*CCR5-Δ32/Δ32*) are almost completely resistant to HIV-1 infections, even among individuals who are at high risk for infection. The mechanism of *CCR2* genetic restriction is less obvious, but likely relates to the ability of CCR2 also to serve as a receptor for HIV-1 infection of macrophages, monocytes and T-cells (Schuitamaker *et al., J. Virol.* 64:356, 1991; Schuitemaker *et al. J. Virol* 66:1354, 1992; Asjo, *Lancet* ii:660, 1986; Connor *et al., J. Virol.* 67:1772, 1993; Roos *et al., J. Infect. Dis.* 165:427, 1992; Zhu *et al., Science* 261:1179, 1993).

HIV-1 strains isolated from recently infected individuals are predominantly M-tropic (macrophage or monocyte lineage tropic), NSI (non-syncytium inducing), and co-opt CC-chemokine receptor proteins as entry ports in combination with CD4 molecules. Later in the course of HIV-1 infection, near the time at which AIDS symptoms are observed, a preponderance of T-tropic (T-lymphocyte cell line tropic) strains have been recovered . T-tropic strains induce the formation of syncytia in CD4-positive cell lines, infect PBMCs faster, and replicate more aggressively than the early M-tropic isolates. The occurrence of T-tropic isolates usually precedes a precipitous drop in CD4 T-cells suggesting that these viruses may contribute to T-cell depletion. T-tropic HIV-1 enters target cells using both CD4 and CXCR4 as obligate co-receptors. The term CXCR4 is preferred, however, the terms fusin or HFAF have also been used to refer to the same molecule. Comparison of the nucleotide sequence of the cDNA encoding CXCR4 against a computer database revealed that CXCR4 is a member of the 7-transmembrane segment superfamily of G-protein-coupled cell surface molecules.

Stromal derived factor (SDF-1, also called pre-B-cell growth stimulating factor), a powerful chemoattractant cytokine, is the natural ligand for CXCR4 and recent

experiments have shown SDF-1α (one of two transcriptional splice variants of the *SDF1* gene) is capable of down regulating CXCR4 on cells by induction of endocytosis, effectively blocking infection by T-tropic but not M-tropic HIV-1 strains.

Without an effective diagnostic or prognostic test or effective vaccine, the number of individuals infected with HIV will likely increase substantially. Furthermore, in the absence of effective therapy, most individuals infected with HIV will develop acquired immune deficiency syndrome (AIDS) and succumb to either opportunistic infections and malignancies that result from the deterioration of the immune system, or the direct pathogenic effects of the virus. Despite the present availability of some anti-HIV agents that slow disease progression, a pressing need remains for more effective therapeutics and drug combinations. To date, there has been no report of any anti-HIV therapeutic that relates to CXCR4 or its ligand, SDF-1.

Summary of the Invention

The present invention is based on the discovery of a correlation between the presence of a mutation at nucleotide position 801 of SDF-1 and resistance to HIV infection. Based on this discovery, it is an object of the present invention to provide diagnostic and therapeutic approaches for identifying the mutation and down-regulating the CXCR4 receptor, respectively.

In a first embodiment, the invention provides an isolated polynucleotide encoding a stromal cell derived factor-1 (SDF-1) variant having a nucleotide sequence set forth in SEQ ID NO:1. SEQ ID NO:1 shows the nucleotide sequence of wild-type SDF-1, with a G to A transition mutation at position 801.

In another embodiment, the invention provides a method for determining the prognosis of a subject exposed to HIV-1. The method is based on determining the presence of a SDF-1 variant nucleic acid in cells of the subject and correlating the

presence of the variant on both alleles with prognosis of said subject. The SDF-1 variant described in the present invention is a recessive mutation, thus both alleles must exhibit the mutation to affect HIV-1 susceptibility. The cells are preferably PBLs from the subject. Preferably the subject is a human.

In yet another embodiment, the invention provides a method of determining susceptibility of a subject to HIV infection by determining the SDF-1 allelic profile of a subject. The method includes isolating the SDF-1 nucleic acid sequence and determining the presence or absence of a mutation in SDF-1 nucleic acid.

The invention also provides a method of inhibiting membrane fusion between HIV and a target cell that expresses CXCR4 or between an HIV-infected cell and a CD4 positive uninfected cell that expresses CXCR4, including contacting the target or CD4/CXCR4 positive cell with a CXCR4 down-regulating effective amount of a SDF-1 variant, thereby inhibiting membrane fusion. The contacting may be by *in vivo* administration to a subject or by *ex vivo* administration to a cell, for example.

In another embodiment, the invention provides a method of treating a subject having or at risk of having an HIV infection or disorder by administering to the subject, a therapeutically effective amount of an SDF-1 variant, such as SEQ ID NO:1. The subject treated by the method of the invention may be suffering from AIDS or ARC.

In yet another embodiment, the invention provides a method of treating a subject having a disorder associated with expression of CXCR4 including administering to the subject, an SDF-1 variant that suppresses CXCR4.

The subject of the invention is well suited for preparation of a kit for determining the SDF-1 allelic profile of a subject. The kit includes amplification primers or hybridization probes which detect a transition mutation of G to A at nucleotide 801.

Brief Description of the Drawing

Figure 1A-I are graphs showing Kaplan-Meier survival curves demonstrating the effect of the SDF1-3'A/3'A genotype on progression to AIDS-1993, AIDS-1987, and death in: the MACS cohort (panels A-C) ; Caucasians in the ALIVE, MACS, MHCS, SFCC combined cohorts (panels D-F) ; and all ethnic groups in the four combined cohorts (panels G-I) .

Figure 2A-F are bar graphs which define disease category analysis of SDF1-3'A allele (panels A-C) and genotype (panels D-F) frequencies for each cohort and combined cohorts for the three endpoints, AIDS-1993 (panels A and D), AIDS-1987 (panels B and E), and death (panels C and F). Cutoffs, in years, were chosen as the time approximately half of all seroconverters had progressed to the outcomes. Times for the cutoffs were: 1) AIDS-1993; 7.5 year; 2) AIDS-1987, 8.5 years; 3) Death, 9.5 years.

Figure 3A-I are graphs of Kaplan-Meier survival curves for the four protective genotypes for SDF1, CCR2, and CCR5 versus wild-type [+/+] at the three loci. Criteria for patient inclusion are as described in Fig. 1A-I. The protective genotypes are: SDF1-3'A/3'A; CCR2-[+/64I], [64I/64I]; and CCR5-+/Δ32. The four curves represent the following genotypes; 1) blue-+/+ at SDF1, CCR2 and CCR5; 2) green-CCR2/5 protein: one or more CCR2/5 protective genotypes and SDF-+/+ ; 3) orange-SDF1: SDF1-3'A/3'A and CCR2/5-+/+; 4) pink-SDF1 and CCR2/5: SDF1-3'A/3'A and protection by one or more CCR2/5 protective genotype versus +/+. n=number of individuals; p=log likelihood p value; and RH=relative hazard based on the Cox proportional hazards model (Center for Disease Control, *Morb. Mort. Wkly. Rep.* 41 (18 December 1992)). x-indicates single events; • indicate patient censoring. Summary statistics for each cohort and the combined cohort analyses are presented in Table 2.

Figure 4A-C shows the frequencies of the protective SDF1-3'A/3'A genotype alone (black) or in combination with at least one CCR2/5 protective genotype (CCR5-+/Δ32, CCR2-+/64I, and CCR2-64I/64I, cross hatch) in six intervals of increasing survivorship from midpoint (seroconverters) or imputed (seroprevalents) seroconversion dates in Caucasians. Genotypic frequencies were determined separately for time to AIDS-1993 (panel A), AIDS-1987 (panel B), and to death (panel C) using seroconverters progressing to the outcome in less than 3.5 years, and including seroconverters and seroprevalents progressing to the outcomes in 3.5 < 7 years, < 10 years, 10 < 13 years, and 13 < 16 years, and ≥ 16 years.

Figure 5 shows the nucleotide sequence of SDF1-3'A (SEQ ID NO:1), which is the SDF-1 variant of the invention.

Other objects, features and advantages of the present invention will become apparent from the following detailed description. It should be understood, however, that the detailed description and the specific examples, while indicating preferred embodiments of the invention, are given by way of illustration only, since various changes and modifications within the spirit and scope of the invention will become apparent to those skilled in the art from this detailed description.

Detailed Description of the Invention

The present invention is based on the identification of a variant of SDF-1 which appears to correlate in some subjects to resistance to HIV-1 infection. This enables therapeutic, prophylactic, prognostic and diagnostic approaches to AIDS. Further, the finding of this variant is useful for therapeutic approaches to inflammatory disorders associated with the expression of CXCR4, the receptor for the wild-type SDF-1 ligand.

Definitions

"Allele" - A gene present in more than one form (different sequence) in a genome, is said to have multiple alleles.

"wt" - Wt stands for the wild type allele of SDF-1, namely the gene without the G to A transition mutation at position 801.

"SDF1-3'A" stands for the mutant allele of SDF-1, found at a frequency of about 0.21 in the caucasian population; 0.162 in the hispanic population; 0.057 in the African American population; and Asians, -0.257. It is also alternatively described as "SDF-1 variant".

"Homozygous" - According to the present invention, the SDF-1 gene (wt or variant) is present, alike most eukaryotic genes, as two copies/genome. If both copies are genetically alike, in regard to the absence or presence of the G to A mutation at position 801, the individual is homozygous, *i.e.*, he is either wt/wt or SDF-3'A/3'A. Since the mutation is recessive, a homozygous mutation only will provide a meaningful protection.

"Heterozygous" -If one copy each of the wt allele and the SDF1-3'A allele are present in one genome, the individual having such a genome is heterozygous. Since the mutation is recessive, a heterozygous mutation is not believed to afford the individual meaningful protection from HIV-1 infection.

"Allelic profile" - A determination of the composition of an individuals genome in regard to the presence or absence, and the copy number, of the SDF1-3'A allele.

Overview

The requirement for available CXCR4 co-receptors during late stage AIDS plus the demonstration that SDF-1 effectively inhibits HIV-1 replication prompted a

polymorphism screen for *SDF-1* structural gene variants that might influence HIV-1 transmission or pathogenesis. 1354 bp of the 3526 bp represented in human SDF-1 transcripts were screened using a series of PCR primers and single strand conformation polymorphism (SSCP/heteroduplex) assays (Kaslow *et al.*, *Am. J. Epidem.* 126:310, 1987; Phair *et al.*, *J. AIDS* 5:490, 1992; Detels *et al.*, *J. Acquir. Immune Defic. Syndr.* 7:1263, 1994) in a group of 144 patients enrolled in epidemiologic cohorts assembled to monitor HIV-1 infection and AIDS. DNA sequence analysis of a common variant revealed a G-A transition mutation at position 801 (counting from the ATG start codon) in the 3' untranslated region of the reference sequence (GenBank L36033). The mutation (designated *SDF1-3'UTR-G801A*, and abbreviated *SDF1-3'A* or SDF-1 variant below and in SEQ ID NO:1) is represented in the *SDF1β* transcripts but not in the *SDF1α* transcript. The variant eliminated an Msp1 site which was converted from SSCP to PCR-RFLP format for rapid screening. The allele and genotype frequency of *SDF1-3'A* was determined in 2860 individuals from the five AIDS cohorts. The cohorts were developed with homosexual, hemophiliac and I.V. drug user groups at risk for HIV-1 infection. The *SDF1-3'A* variant was found in the following allele frequencies: Caucasian-0.21 (N=1833); Hispanics-0.162 (N=130); African American - 0.057 (N=860); and Asians - 0.257 (N=37). Seroconverter patients included 867 subjects with a maximum interval of three years between an HIV-1 antibody negative test date and their first HIV-1 antibody positive test date. Seroconversion date was the mid-point between the last HIV-1 antibody negative and first positive clinic visits. Ninety patients enrolled in the SFCC study before December 31, 1980 were included using imputed seroconversion dates based on their date of first HIV-1 antibody positive test, because the likelihood of infection before January 1, 1978 (a 3 year window of infection) was extremely low (≤0.01). Seroconversion dates for the imputed SFCC subjects were set at 60 days, 120 days and 180 days before the date for first antibody positive visit for patients enrolled in 1978, 1979 and 1980 respectively.

Isolated Nucleic Acid Sequences for SDF1-3'A

In one embodiment, the invention provides an isolated polynucleotide encoding a stromal cell derived factor (SDF-1) variant (SDF1-3'A) having a nucleotide sequence which differs from wild-type SDF-1 by a single mutation. Specifically, the variant of the invention, having a G to A transition at nucleotide 801 (counting from the ATG start codon) is set forth in Figure 5 and SEQ ID NO:1. A "variant" as used herein refers to a nucleotide sequence that is altered as compared to the wild-type sequence. An exemplary variant of the invention differs from wild-type SDF-1 by only one nucleotide although other nucleotide changes are also included as long as the SDF-1 allele still correlates with a decreased progression to AIDS and preferably is still is able to down-regulate the CXCR4 receptor.

The term "isolated" as used herein includes polynucleotides substantially free of other nucleic acids, proteins, lipids, carbohydrates or other materials with which it is naturally associated. Polynucleotide sequences of the invention include DNA and RNA sequences which encode SDF-1 and which have a G to A transition mutation at nucleotide 801. It is understood that all polynucleotides encoding all or a portion of SDF1-3'A, but which include nucleotide 801, are also included herein, as long as they encode a polypeptide with SDF-1 activity (*e.g.*, bind to and down-regulate CXCR4). Such polynucleotides include naturally occurring, synthetic, and intentionally manipulated polynucleotides. As an example, wild-type SDF-1 polynucleotide may be subjected to site-directed mutagenesis to produce a G to A substitution at position 801. The polynucleotides of the invention include sequences that are degenerate as a result of the genetic code. There are 20 natural amino acids, most of which are specified by more than one codon. Therefore, all degenerate nucleotide sequences are included in the invention as long as the amino acid sequence of SDF-1 polypeptide encoded by the nucleotide sequence is functionally unchanged (e.g., down-regulates CXCR4).

The polynucleotide encoding SDF1-3'A includes the nucleotide sequence in FIGURE 5 (SEQ ID NO:1), as well as nucleic acid sequences complementary to that sequence. A complementary sequence may include an antisense nucleotide. When the sequence is RNA, the deoxyribonucleotides A, G, C, and T of FIGURE 5 are replaced by ribonucleotides A, G, C, and U, respectively. Also included in the invention are fragments (portions) of the above-described nucleic acid sequences that are at least 15 bases in length, which is sufficient to permit the fragment to selectively hybridize to DNA of the variant SDF1-3'A. The fragments of the invention encompass position 801 (*e.g.*, SEQ ID NO:1). "Selective hybridization" as used herein refers to hybridization under moderately stringent or highly stringent physiological conditions (See, for example, the techniques described in Maniatis *et al.*, 1989 Molecular Cloning A Laboratory Manual, Cold Spring Harbor Laboratory, N.Y., incorporated herein by reference), which distinguishes related from unrelated nucleotide sequences.

In nucleic acid hybridization reactions, the conditions used to achieve a particular level of stringency will vary, depending on the nature of the nucleic acids being hybridized. For example, the length, degree of complementarity, nucleotide sequence composition (*e.g.*, GC v. AT content), and nucleic acid type (*e.g.*, RNA v. DNA) of the hybridizing regions of the nucleic acids can be considered in selecting hybridization conditions. An additional consideration is whether one of the nucleic acids is immobilized, for example, on a filter.

An example of progressively higher stringency conditions is as follows: 2 x SSC/0.1% SDS at about room temperature (hybridization conditions); 0.2 x SSC/0.1% SDS at about room temperature (low stringency conditions); 0.2 x SSC/0.1% SDS at about 42°C (moderate stringency conditions); and 0.1 x SSC at about 68°C (high stringency conditions). Washing can be carried out using only one of these conditions, *e.g.*, high stringency conditions, or each of the conditions can be used, *e.g.*, for 10-15 minutes each, in the order listed above, repeating any or all of the

steps listed. However, as mentioned above, optimal conditions will vary, depending on the particular hybridization reaction involved, and can be determined empirically.

DNA sequences of the invention can be obtained by several methods. For example, the DNA can be isolated using hybridization or computer-based techniques which are well known in the art. These include, but are not limited to: 1) hybridization of genomic or cDNA libraries with probes to detect homologous nucleotide sequences; 2) antibody screening of expression libraries to detect cloned DNA fragments with shared structural features; 3) polymerase chain reaction (PCR) on genomic DNA or cDNA using primers capable of annealing to the DNA sequence of interest; 4) computer searches of sequence databases for similar sequences; and 5) differential screening of a subtracted DNA library.

Preferably the SDF1-3'A polynucleotide of the invention is derived from a mammalian organism. Screening procedures which rely on nucleic acid hybridization make it possible to isolate any gene sequence from any organism, provided the appropriate probe is available. Oligonucleotide probes, which correspond to a part of the sequence encoding the protein in question, can be synthesized chemically. This requires that short, oligopeptide stretches of amino acid sequence must be known. The DNA sequence encoding the protein can be deduced from the genetic code, however, the degeneracy of the code must be taken into account. It is possible to perform a mixed addition reaction when the sequence is degenerate. This includes a heterogeneous mixture of denatured double-stranded DNA. For such screening, hybridization is preferably performed on either single-stranded DNA or denatured double-stranded DNA. Hybridization is particularly useful in the detection of cDNA clones derived from sources where an extremely low amount of mRNA sequences relating to the polypeptide of interest are present. In other words, by using stringent hybridization conditions directed to avoid non-specific binding, it is possible, for example, to allow the autoradiographic visualization of a specific cDNA clone by the

hybridization of the target DNA to that single probe in the mixture which is its complete complement (Wallace, *et al., Nucl. Acid Res.*, 9:879, 1981).

When the entire sequence of amino acid residues of the desired polypeptide is not known, the direct synthesis of DNA sequences is not possible and the method of choice is the synthesis of cDNA sequences. Among the standard procedures for isolating cDNA sequences of interest is the formation of plasmid- or phage-carrying cDNA libraries which are derived from reverse transcription of mRNA which is abundant in donor cells that have a high level of genetic expression. When used in combination with polymerase chain reaction technology, even rare expression products can be cloned. In those cases where significant portions of the amino acid sequence of the polypeptide are known, the production of labeled single or double-stranded DNA or RNA probe sequences duplicating a sequence putatively present in the target cDNA may be employed in DNA/DNA hybridization procedures which are carried out on cloned copies of the cDNA which have been denatured into a single-stranded form (Jay, *et al., Nucl. Acid Res.*, 11:2325, 1983).

Other alterations in SDF1-3'A nucleic acid include intragenic mutations (*e.g.*, point mutation, nonsense (stop), missense, splice site and frameshift) and heterozygous or homozygous deletions. Detection of such alterations can be done by standard methods known to those of skill in the art including sequence analysis, Southern blot analysis, PCR based analyses (*e.g.*, multiplex PCR, sequence tagged sites (STSs)) and *in situ* hybridization. Such proteins can be analyzed by standard SDS-PAGE and/or immuno-precipitation analysis and/or Western blot analysis, for example.

DNA sequences encoding SDF1-3'A can be expressed *in vitro* by DNA transfer into a suitable host cell. "Host cells" are cells in which a vector can be propagated and its DNA expressed. The term also includes any progeny of the subject host cell. It is understood that all progeny may not be identical to the parental cell since there may be mutations that occur during replication. However, such progeny are included

when the term "host cell" is used. Methods of stable transfer, meaning that the foreign DNA is continuously maintained in the host, are known in the art.

In the present invention, the SDF1-3'A polynucleotide sequences may be inserted into a recombinant expression vector. The term "recombinant expression vector" refers to a plasmid, virus or other vehicle known in the art that has been manipulated by insertion or incorporation of the SDF1-3'A genetic sequences. Such expression vectors contain a promoter sequence which facilitates the efficient transcription of the inserted genetic sequence of the host. The expression vector typically contains an origin of replication, a promoter, as well as specific genes which allow phenotypic selection of the transformed cells. Vectors suitable for use in the present invention include, but are not limited to the T7-based expression vector for expression in bacteria (Rosenberg, *et al.*, *Gene* ,56:125, 1987), the pMSXND expression vector for expression in mammalian cells (Lee and Nathans, *J. Biol. Chem.*, 263:3521, 1988) and baculovirus-derived vectors for expression in insect cells. The DNA segment can be present in the vector operably linked to regulatory elements, for example, a promoter (*e.g.*, T7, metallothionein I, or polyhedrin promoters).

Polynucleotide sequences encoding SDF1-3'A can be expressed in either prokaryotes or eukaryotes. Hosts can include microbial, yeast, insect and mammalian organisms. Methods of expressing DNA sequences having eukaryotic or viral sequences in prokaryotes are well known in the art. Biologically functional viral and plasmid DNA vectors capable of expression and replication in a host are known in the art. Such vectors are used to incorporate DNA sequences of the invention.

Methods which are well known to those skilled in the art can be used to construct expression vectors containing the SDF1-3'A coding sequence and appropriate transcriptional/translational control signals. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* recombination/genetic techniques. (See, for example, the techniques described in

Maniatis *et al.*, 1989 Molecular Cloning A Laboratory Manual, Cold Spring Harbor Laboratory, N.Y.)

A variety of host-expression vector systems may be utilized to express the SDF1-3'A coding sequence. These include but are not limited to microorganisms such as bacteria transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing the SDF1-3'A coding sequence; yeast transformed with recombinant yeast expression vectors containing the SDF1-3'A coding sequence; plant cell systems infected with recombinant virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (*e.g.*, Ti plasmid) containing the SDF1-3'A coding sequence; insect cell systems infected with recombinant virus expression vectors (*e.g.*, baculovirus) containing the SDF1-3'A coding sequence; or animal cell systems infected with recombinant virus expression vectors (*e.g.*, retroviruses, adenovirus, vaccinia virus) containing the SDF1-3'A coding sequence, or transformed animal cell systems engineered for stable expression. Since SDF1-3'A has not been confirmed to contain carbohydrates, both bacterial expression systems as well as those that provide for translational and post-translational modifications may be used; *e.g.*, mammalian, insect, yeast or plant expression systems.

Depending on the host/vector system utilized, any of a number of suitable transcription and translation elements, including constitutive and inducible promoters, transcription enhancer elements, transcription terminators, etc. may be used in the expression vector (see *e.g.*, Bitter *et al.*, 1987, Methods in Enzymology 153:516-544). For example, when cloning in bacterial systems, inducible promoters such as pL of bacteriophage γ, plac, ptrp, ptac (ptrp-lac hybrid promoter) and the like may be used. When cloning in mammalian cell systems, promoters derived from the genome of mammalian cells (*e.g.*, metallothionein promoter) or from mammalian viruses (*e.g.*, the retrovirus long terminal repeat; the adenovirus late promoter; the vaccinia virus

7.5K promoter) may be used. Promoters produced by recombinant DNA or synthetic techniques may also be used to provide for transcription of the inserted SDF1-3'A coding sequence.

In yeast, a number of vectors containing constitutive or inducible promoters may be used. For a review see, Current Protocols in Molecular Biology, Vol. 2, 1988, Ed. Ausubel *et al.*, Greens Publish. Assoc. & Wiley Interscience, Ch. 13; Grant *et al.*, 1987, Expression and Secretion Vectors for Yeast, in Methods in Enzymology, Eds. Wu & Grossman, 31987, Acad. Press, N.Y., Vol. 153, pp.516-544; Glover, 1986, DNA Cloning, Vol. II, IRL Press, Wash., D.C., Ch. 3; and Bitter, 1987, Heterologous Gene Expression in Yeast, Methods in Enzymology, Eds. Berger & Kimmel, Acad. Press, N.Y., Vol. 152, pp. 673-684; and The Molecular Biology of the Yeast Saccharomyces, 1982, Eds. Strathern *et al.*, Cold Spring Harbor Press, Vols. I and II. A constitutive yeast promoter such as ADH or LEU2 or an inducible promoter such as GAL may be used (Cloning in Yeast, Ch. 3, R. Rothstein In: DNA Cloning Vol.11, A Practical Approach, Ed. DM Glover, 1986, IRL Press, Wash., D.C.). Alternatively, vectors may be used which promote integration of foreign DNA sequences into the yeast chromosome.

Eukaryotic systems, and preferably mammalian expression systems, allow for proper post-translational modifications of expressed mammalian proteins to occur. Eukaryotic cells which possess the cellular machinery for proper processing of the primary transcript, glycosylation, phosphorylation, and advantageously, plasma membrane insertion of the gene product may be used as host cells for the expression of SDF1-3'A.

Mammalian cell systems which utilize recombinant viruses or viral elements to direct expression may be engineered. For example, when using adenovirus expression vectors, the SDF1-3'A coding sequence may be ligated to an adenovirus transcription/translation control complex, *e.g.*, the late promoter and tripartite leader

sequence. Alternatively, the vaccinia virus 7.5K promoter may be used. (*e.g.*, see, Mackett *et al.*, 1982, Proc. Natl. Acad. Sci. USA 79: 7415-7419; Mackett *et al.*, 1984, J. Virol. 49: 857-864; Panicali *et al.*, 1982, Proc. Natl. Acad. Sci. USA 79: 4927-4931). Of particular interest are vectors based on bovine papilloma virus which have the ability to replicate as extrachromosomal elements (Sarver, *et al.*, 1981, Mol. Cell. Biol. 1: 486). Shortly after entry of this DNA into mouse cells, the plasmid replicates to about 100 to 200 copies per cell. Transcription of the inserted cDNA does not require integration of the plasmid into the host's chromosome, thereby yielding a high level of expression. These vectors can be used for stable expression by including a selectable marker in the plasmid, such as, for example, the neo gene. Alternatively, the retroviral genome can be modified for use as a vector capable of introducing and directing the expression of the SDF1-3'A gene in host cells (Cone & Mulligan, 1984, *Proc. Natl. Acad. Sci. USA* 81:6349-6353). High level expression may also be achieved using inducible promoters, including, but not limited to, the metallothionine IIA promoter and heat shock promoters.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. Rather than using expression vectors which contain viral origins of replication, host cells can be transformed with the SDF1-3'A cDNA controlled by appropriate expression control elements (*e.g.*, promoter, enhancer, sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn can be cloned and expanded into cell lines. For example, following the introduction of foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. A number of selection systems may be used, including but not limited to the herpes simplex virus thymidine kinase (Wigler, *et al.*, *Cell* 11:223, 1977), hypoxanthine-guanine phosphoribosyl transferase (Szybalska & Szybalski, *Proc. Natl. Acad. Sci. USA* 48:2026, 1962), and adenine phosphoribosyltransferase (Lowy, *et al.*, *Cell* 22: 817,

1980) genes can be employed in tk-, hgprt- or aprt- cells respectively. Also, antimetabolite resistance can be used as the basis of selection for dhfr, which confers resistance to methotrexate (Wigler, *et al.*,, *Natl. Acad. Sci. USA* 77:3567, 1980; O'Hare, *et al.*, *Proc. Natl. Acad. Sci. USA* 78:1527, 1981); gpt, which confers resistance to mycophenolic acid (Mulligan & Berg, *Proc. Natl. Acad. Sci. USA* 78: 2072, 1981; neo, which confers resistance to the aminoglycoside G-418 (Colberre-Garapin, *et al.*, *J. Mol. Biol.* 150:1, 1981); and hygro, which confers resistance to hygromycin (Santerre, *et al.*, *Gene* 30:147, 1984) genes. Recently, additional selectable genes have been described, namely trpB, which allows cells to utilize indole in place of tryptophan; hisD, which allows cells to utilize histinol in place of histidine (Hartman & Mulligan, *Proc. Natl. Acad. Sci. USA* 85:8047, 1988); and ODC (ornithine decarboxylase) which confers resistance to the ornithine decarboxylase inhibitor, 2-(difluoromethyl)-DL-ornithine, DFMO (McConlogue, In: Current Communications in Molecular Biology, Cold Spring Harbor Laboratory ed., 1987).

When the host is a eukaryote, such methods of transfection of DNA as calcium phosphate co-precipitates, conventional mechanical procedures such as microinjection, electroporation, insertion of a plasmid encased in liposomes, or virus vectors may be used. Eukaryotic cells can also be cotransformed with DNA sequences encoding the SDF1-3'A of the invention, and a second foreign DNA molecule encoding a selectable phenotype, such as the herpes simplex thymidine kinase gene. Another method is to use a eukaryotic viral vector, such as simian virus 40 (SV40) or bovine papilloma virus, to transiently infect or transform eukaryotic cells and express the protein. (see for example, *Eukaryotic Viral Vectors*, Cold Spring Harbor Laboratory, Gluzman ed., 1982).

Methods for Prognosis

In another embodiment, the invention includes a method for determining the prognosis of a patient exposed to HIV-1. The patient may be asymptomatic or symptomatic for infection with HIV-1. The prognosis of the patient is ascertained by

determining the presence of a SDF-1 variant nucleic acid in cells of the subject and correlating the presence of the variant on both alleles with prognosis of the subject. Appropriate cells include those susceptible to infection by HIV-1, such as peripheral blood leukocytes. Identification of the presence of a SDF1-3'A variant present on both alleles can be determined routinely by PCR amplification of SDF-1 as described herein followed by sequence analysis.

The allelic profile of a patient can be determined by employment of PCR technology. The target nucleic acid to be amplified by PCR would be either the SDF-1 RNA (formation of a cDNA) or, in a preferred embodiment, the SDF-1 gene. Primers would be designed on sequences flanking the putative SDF1-3'A mutation site. By judiciously choosing the primers one can obtain a fragment whose size is indicative of the presence or absence of the deletion. For example, fragments whose size are between 75 to 450 nucleotides are preferred, though PCR products shorter and longer are acceptable. The size consideration relates mostly to the ability to visualize the product after separation on an agarose gel. One skilled in the art would recognize many variations on this motif. For example, the PCR reaction may contain labeled oligonucleotides to facilitate subsequent detection of the PCR product. The label can be, for example, radiolabeled nucleotides, or biotin incorporating nucleotides. Another variation of the technique would employ slurries other than agarose, or filters, for size separation of the PCR product. According to such a diagnostic procedure employing PCR, either wt or SDF1-3'A homozygote will produce one product of a discrete, expected size, while the heterozygous individual will be identified by production of two differently sized products.

More specifically, using a series of PCR primers and single strand conformation polymorphism (SSCP/heteroduplex) assays followed by sequence analysis, detection of SDF1-3'A can be accomplished. Primers used in the present invention were as follows:

5'UTR: (197 bp)	GGC AGG TGG CGA GCT TGA GC (F)	(SEQ ID NO:2) and
	CTG GAG GGC CGC TTA TTG TC (R)	(SEQ ID NO:3);
Exon l: (130 bp)	AGC CGC ATT GCC CGC TCG GCG TC (F)	(SEQ ID NO:4)
and	CGT CGC TGA GGC AGA GCG CGG TC (R)	(SEQ ID NO:5);
Exon 2: (218 bp)	CNA AAT CTG NCA GGG TAG TA (F)	(SEQ ID NO:6) and
	TCG TTA GAT GCA ACT ATG TTC (R)	(SEQ ID NO:7);
Exon 3: (189 bp)	AGC CGC GCT TTC CTC CTG TGC (F)	(SEQ ID NO:8) and
	TAG TTT TCC TCG AGT GGG TC (R)	(SEQ ID NO:9);
Exon 4: (318 bp)	CTG TCC TGC TGG AGC TGG C (F)	(SEQ ID NO:10) and
	TTT CAG AGC TGG GCT CCT AC (R)	(SEQ ID NO:11);
3'UTR: (302 bp)	CAG TCA ACC TGG GCA AAG CC (F)	(SEQ ID NO:12) and
	AGC TTT GGT CCT GAG AGT CC (R)	(SEQ ID NO:13).

Generally, the prognosis of the patient improves by the detection of SDF1-3'A on both alleles, since the mutation is recessive. Further, it may be desirable to analyze a mutation in the CCR2 receptor and/or analyze a mutation in the CCR5 receptor. Common alleles within the coding region for the chemokine and M-tropic HIV-1 co-receptor genes, CCR5 and CCR2, have been shown to delay the rate of progression to AIDS (4-9). The mutant alleles CCR5-Δ32 and CCR2-64I are dominant, genetically independent, and equally protective. An estimated 25-30% of long-term survivors who remain AIDS-free for > 16 years can be attributed to a protective genotype for either CCR5-Δ32 or CCR2-64I (6,9). A survival analysis of the relative contributions of CCR5-Δ32, CCR2-64I, and SDF1-3'A genotypes (Fig. 3, Table 2) reaffirm the protective effects of CCR2, CCR5 and SDF1 mutant genotypes on progression to AIDS when the influence of the other protective loci are considered as confounding variables (29,33).

The method for determining prognosis can be used to monitor subjects infected with HIV-1 and can be used to assess the results in clinical trials for pharmaceuticals, vaccines and other therapies as well.

Methods for Determining Susceptibility

In yet another embodiment, the invention provides a method of determining susceptibility of a subject to HIV infection by determining the SDF-1 allelic profile of a subject. The method includes isolating the SDF-1 nucleic acid sequence and determining the presence or absence of a mutation in SDF-1 nucleic acid similar to the method described above. AIDS and infection by HIV has been recognized primarily in "at risk" groups, including homosexually active males, intravenous drug users, recipients of blood or blood products, and certain populations from Central Africa and the Caribbean. The AIDS syndrome has also been recognized in heterosexual partners of individuals in all "at risk" groups and in infants of affected mothers. Thus, these groups of individual, in addition to other lower risk groups would be ideally suited for practicing the method of the invention to determine susceptibility.

The method of the invention is typically performed as described above, by PCR amplification of nucleic acid containing SDF-1 as described herein followed by sequence analysis. As in the method above, it may be desirable to determine the CCR2 and/or CCR5 allelic profile of the subject, in addition to the SDF-1 allelic profile. As described above and in the Examples below, it is believed that the SDF-1 variant and the CCR2 and CCR5 mutations act in a protective manner and the effect is additive.

Determination of the likelihood of an initial HIV infection to be sustained in the organism, to led to AIDS, and the likely speed of the disease's development leads to more rational choices of treatment. Diagnostics are enabled by the present invention in that it is recognized that the infection is inhibited or reduced by individuals homozygous in regard to the SDF1-3'A allele and disease progression is reduced in individuals also having a CCR2 and/or CCR5 mutation. The amplification of a nucleic acid can be accomplished by one of a number of methods known to one skilled in the art. By way of example, amplification by PCR is described below.

Method of Inhibiting HIV Infection

The invention also provides a method of inhibiting membrane fusion between HIV and a target cell that expresses CXCR4 or between an HIV-infected cell and a CD4 positive uninfected cell that expresses CXCR4, including contacting the target or CD4/CXCR4 positive cell with a CXCR4 down-regulating effective amount of a SDF-1 variant, thereby inhibiting membrane fusion. Preferably, the SDF-1 variant is SDF1-3'A (SEQ ID NO:1) as described herein. The contacting may be by *in vivo* administration to a subject or by *ex vivo* administration to a cell, for example. While not wanting to be bound by a particular theory, it is believed that down-regulation by SDF1-3'A of the invention blocks the emergence and spread of T-tropic HIV-1, which requires CXCR4 as a co-receptor with CD4 to other cells.

Methods of Treatment and Delivery of Nucleic Acids to Cells

In yet another embodiment, the invention provides a method of treating a subject having a disorder associated with expression of CXCR4 including administering to the subject, an SDF-1 variant that suppresses CXCR4. Such disorders include not only disorders associated with HIV-1 infection, but also inflammatory disorders.

Administration of an SDF1-3'A polynucleotide to a subject, either as a naked, synthetic polynucleotide or as part of an expression vector, can be effected via any common route (oral, nasal, buccal, rectal, vaginal, or topical), or by subcutaneous, intramuscular, intra-peritoneal, or intravenous injection. Pharmaceutical compositions of the present invention, however, are advantageously administered in the form of injectable compositions. A typical composition for such purpose comprises a pharmaceutically acceptable solvent or diluent and other suitable, physiologic compounds. For instance, the composition may contain polynucleotide and about 10 mg of human serum albumin per milliliter of a phosphate buffer containing NaCl.

As much as 700 milligrams of SDF1-3'A polynucleotide has been administered intravenously to a patient over a course of 10 days (*i.e.*, 0.05 mg/kg/hour) without signs of toxicity. Sterling, "Systemic Antisense Treatment Reported," *Genetic Engineering News* 12:1, 28 (1992). The toxicity of a specific oligonucleotide at various levels can be estimated in animal tests by criteria and methods well known to one skilled in the art.

SDF1-3'A polynucleotides expression vectors can be encapsulated within liposomes using standard techniques. Cationic liposomes would be preferred for delivery of nucleic acids. A variety of different liposome compositions and methods for synthesis are known to those of skill in the art. See, for example, U.S. Patent No. 4,844,904, No. 5,000,959, No. 4,863,740 and No. 4,975,282, the respective contents of which are hereby incorporated by reference.

Liposomes are sometimes targeted to the cell type or tissue of interest (here PBLs or CD4+ cells) by the addition to the liposome preparation of a ligand, usually a polypeptide, for which a corresponding cellular receptor has been identified. In this case, a likely such ligand would be gp120 or gp21, or fragments thereof. Examples of a cell receptors previously targeted include folate receptor which has recently been identified as a prominent tumor marker, especially in ovarian carcinomas. KB cells are known to vastly overexpress the folate receptor. See Campbell *et al.*, *Cancer Res.* 51:6125 (1991). Yet other targeting ligands have been examined for liposome targeting including transferrin, protein A, ApoE, P-glycoprotein, α_2-macroglobin, insulin, asiolofetuin, asialoorosomucoid, monoclonal antibodies with a variety of tissue specificity, biotin, galactose or lactose containing haptens (monovalent and tri-antennary), mannose, dinitrophenol, and vitamin B_{12}. The ligands are covalently conjugated to a lipid anchor in either pre-formed liposomes or are incorporated during liposome preparation. See Lee & Low, *J. Biol. Chem.* 269:3198 (1994); *Biochim. Biophys. Acta* 1233:134 (1995).

In one embodiment of the invention, the association of the SDF1-3'A polypeptide of the invention with an agent as described above includes association with additional targeting agents. For example, in order to gain access to the cytoplasm, a nucleic acid based therapeutic must overcome the plasma membrane barrier. In nature, viral fusion peptides facilitate the delivery of viral DNA into the cytoplasm by promoting viral membrane fusion with the plasma membrane. For recent reviews on this subject, see Stegmann *et al.*, *Ann. Rev. Biophys. Chem.* 18:187 (1989). For the influenza virus, the hemagglutinin (trimer) HA peptide N-terminal segment (a hydrophobic helical sequence) is exposed due to a conformational change induced by acidic pH in the endosome (pH 5-6), inserts into the target membrane, and mediates the fusion between the virus and the target endosomal membrane. See Weber *et al.*, *J. Biol. Chem.* 269:18353(1994). Recently, several amphipathic helix-forming oligopeptides have been designed to imitate the behavior of the viral fusion peptide. See, for example, Haensler & Szoka, *Biocon. Chem.* 4:372 (1993).

Nuclear localization signal peptides, when attached covalently to a macromolecule such as a protein, have been shown to facilitate their translocation into the nucleus. See Goldfarb *et al.*, *Nature* 322:641 (1986); Shreiber *et al.*, *Med. Sci.* 8:134-39 (1992). By the combination of cellular targeting by the polypeptide of the invention and nucleus targeting, yet other agents could be delivered in the nucleus of specific cells, for example DNA molecules.

According to embodiments of the invention, treatment of prevention of HIV-1 infection is achieved by introduction into a patient of CD4+ cells or bone marrow cells derived from donors. Ideally, such donors would be homozygous in respect to SDF1-3'A. The data suggests that about 21 in 100 caucasian individuals would be homozygous for the deletion allele.

Donors will furthermore be HLA-matched individuals. They preferably would be blood relatives. The blood typing of a potential donor in terms of compatibility is

well known to one skilled in the art. See, for example, Beatty *et al., Transplantation*, 45: 714-8 (1988).

Purification of CD4+ cells from lymphocytes is also well known to one skilled in the art. Generally, such purification involves use of an antibody directed to the CD4 epitope.

In another embodiment of the invention, umbilical cord blood stem cells are employed for transplantation. In yet another embodiment, CD34+ stem cells are isolated (usually by use of antibodies) and used for transplantation. Cord blood stem cells and CD34+ cells are better tolerated, *i.e.* host rejection is limited, when compared with rejection of bone marrow or CD4+ cells. CD34+ and cord cell transplantation can be used for adult recipients. Furthermore, they are a preferred source of transplantation tissue for infants.

In yet another embodiment of the invention, bone marrow cells or CD4+ blood cells are isolated from the patient himself. The isolated cells are transfected with a vector engineered to express an SDF1-3'A oligonucleotide, and transfected cells are selected. Vectors for transfection of eukaryotic cells are well known in the art. Such vectors have an origin of replication which allows replication and maintenance in the transfected cell. The origin of replication may be a viral origin of replication. One often used viral origin of replication is the SV40 replication region. Furthermore, the vector, to be useful, should contain a marker so transfected cells can be selected. Such a marker often is a drug resistance gene. For example, the *neo* gene confirming the resistance to G418 is often used.

It is very important to realize that all of the transplantation treatments described above are more likely to be successful in conjunction with more typical viral or anti HIV treatments. In a preferred embodiment, transplantation as described above would be accompanied by on-going antiviral treatments and more specifically anti-HIV-1

treatments. Furthermore standard transplantation methodologies would generally be employed, which may contain such additional treatments as temporary immune suppression or lymphocyte growth and cytokine stimulation. Indeed, all method of treatments embodied by the present invention, the SDF1-3'A approach as described above or the presently described transplantation or antibody treatment would each benefit and are compatible with standard anti-HIV-1 treatments. In a preferred embodiment, any treatment of the invention is augmented by known anti-HIV treatments. Such anti-HIV-1 treatments known to date include use of modified oligonucleotides, use of specific proteases, and specific anti-viral RNA nucleases.

Delivery of the Therapeutic Agents to a Patient

Delivery of any of the above therapeutic agents, including SDF1-3'A oligonucleotides, or isolated cells, would require administration to the patient of therapeutically effective doses. "Administering" the pharmaceutical composition of the present invention may be accomplished by any means known to the skilled artisan.

The pharmaceutical compositions are preferably prepared and administered in dose units. Solid dose units are tablets, capsules and suppositories. For treatment of a patient, depending on activity of the compound, manner of administration, nature and severity of the disorder, age and body weight of the patient, different daily doses are necessary. Under certain circumstances, however, higher or lower daily doses may be appropriate. The administration of the daily dose can be carried out both by single administration in the form of an individual dose unit or else several smaller dose units and also by multiple administration of subdivided doses at specific intervals.

The pharmaceutical compositions according to the invention are preferably administered intravenously. However, other routes of administration is within the scope of the inventor. Thus, the pharmaceutical compositions can be administered topically, intravenously, orally or parenterally or as implants, but even rectal use is

possible in principle. Suitable solid or liquid pharmaceutical preparation forms are, for example, granules, powders, tablets, coated tablets, (micro)capsules, suppositories, syrups, emulsions, suspensions, creams, aerosols, drops or injectable solution in ampule form and also preparations with protracted release of active compounds, in whose preparation excipients and additives and/or auxiliaries such as disintegrants, binders, coating agents, swelling agents, lubricants, flavorings, sweeteners or solubilizers are customarily used as described above. The pharmaceutical compositions are suitable for use in a variety of drug delivery systems. For a brief review of present methods for drug delivery, see Langer, *Science*, 249:1527(1990), which is incorporated herein by reference.

The pharmaceutical compositions according to the invention may be administered locally or systemically. By "therapeutically effective dose" is meant the quantity of a compound according to the invention necessary to prevent, to cure or at least partially arrest the symptoms of the disease and its complications. Amounts effective for this use will, of course, depend on the severity of the disease and the weight and general state of the patient. Typically, dosages used *in vitro* may provide useful guidance in the amounts useful for *in situ* administration of the pharmaceutical composition, and animal models may be used to determine effective dosages for treatment of particular disorders. Various considerations are described, *e.g.*, in Gilman *et al.* (eds.) (1990) GOODMAN AND GILMAN'S: THE PHARMACOLOGICAL BASES OF THERAPEUTICS, 8th ed., Pergamon Press; and REMINGTON'S PHARMACEUTICAL SCIENCES, 17th ed. (1990), Mack Publishing Co., Easton, Pa., each of which is herein incorporated by reference.

Methods of Treatment

In another embodiment, the invention provides a method of treating a subject having or at risk of having an HIV infection or disorder by administering to the subject, a therapeutically effective amount of an SDF-1 variant, such as SEQ ID NO:1. The

subject treated by the method of the invention may be suffering from AIDS or ARC. Administration of the variant may be by standard techniques as described above. The SDF-1 variant is introduced into the cell using a carrier, such as a vector. Administration can be *in vivo* or *ex vivo*.

Introduction of SDF1-3'A nucleic acids into cells affected by a SDF-1 disorder, for the purpose of gene therapy, can be achieved using a recombinant expression vector, such as a chimeric virus or a colloidal dispersion system, such as a targeted liposome (see above discussion). Those of skill in this art know or can easily ascertain the appropriate route and means for introduction of SDF1-3'A nucleic acids, without resort to undue experimentation.

Kits for Determinationof Allelic Profiles

The subject of the invention is well suited for preparation of a kit for determining the SDF-1 allelic profile of a subject. The kit includes amplification primers or hybridization probes which detect a transition mutation of G to A at nucleotide 801. Such primers, for example, as described in detail in the Examples, can easily be designed based on the publicly available sequence for SDF-1 (GenBank L36033).

Examples

A role for SDF1-3'A in HIV-1 infection was investigated by genotyping 2419 HIV-1 infected patients and 435 HIV-1 exposed uninfected individuals. No significant differences in SDF1 allele or genotype frequencies were observed in comparisons of exposed (or at risk) uninfected (HIV-1-) vs. infected (HIV-1+) individuals in any of the cohorts. A collection of 138 extremely high risk, exposed-uninfected individuals (those with documented receipt of clotting factor prior to March 1984 when HIV-1 screening commenced or with frequent sexual encounters with high risk partners) also showed SDF1 allele frequencies not significantly different from those of HIV-1-infected individuals. The SDF1 genotype frequencies in each cohort and each HIV-1

infection category conformed to expectations of Hardy-Weinberg equilibrium, further excluding any significant effect of SDF1-3'A on HIV-1 infection.

Example 1

Survival Analysis

To assess the influence of SDF1 genotype on progression of HIV-1 infected patients to AIDS, a survival analyses was performed on a group of 867 seroconverter patients (those whose date of HIV-1 infection could be estimated precisely since they enrolled in the cohort before converting from HIV-1-antibody negative to HIV-1-antibody positive) from four cohorts by comparing the rate of progression to AIDS among different SDF1 genotypes (+/+; +/3'A and 3'A/3'A) using a Cox proportional hazards model. Three AIDS endpoints reflecting advancing morbidity were evaluated: 1) AIDS-1993 definition as stipulated by the CDC (Center for Disease Control, *Morb. Mort. Wkly. Rep.* 36, suppl. 1, August 1987) (*i.e.*, HIV-1 infection plus AIDS-defining illness or decline of CD4 T-lymphocytes to <200 cells/mm^3) or death; 2) the more stringent AIDS-1987 definition or death; and 3) death during follow up for an HIV-1 infected patient (97% of these had AIDS-1993). Additional evidence in support of an increasing appearance of *SDF1-3'A/3'A* protection in late stages of HIV-1 infection involves the identification of a statistically significant difference in the frequency of homozygotes (Fisher's Exact Test: $p \leq 0.01$) in seroconverter (f=3.5%; n=669) versus seroprevalent (f=6.2%; n=743) cohort members. The enrichment of homozygotes in seroprevalent individuals is consistent with late stage protection for two reasons: 1) seroprevalents are seropositive at enrollment and therefore included more patients with long intervals since infection; and 2) certain studies are biased to include more long term survivors than rapid progressors. Two of the cohorts, HGDS and MACS, specifically excluded enrollment of individuals with AIDS defining conditions, and the SFCC selected for long term survivors (HIV-1 infection plus AIDS defining illness). The results of these analyses are illustrated in Fig. 1 and tabulated in Table 1.

Figure 1A-I are graphs showing Kaplan-Meier survival curves demonstrating the effect of the SDF1-3'A/3'A genotype on progression to AIDS-1993, AIDS-1987, and death in: the MACS cohort (panels A-C) ; Caucasians in the ALIVE, MACS, MHCS, SFCC combined cohorts (panels D-F) ; and all ethnic groups in the four combined cohorts (panels G-I) . The analysis was limited to seroconverters with an interval of ≤ 3 years between last seronegative and first seropositive HIV-1 test. The midpoint between these two dates was used to estimate the seroconversion date (Cox, *J.R. Stat Soc. B* 34:187, 1972). The SDF1-3'A/3'A genotype survival was compared to that of SDF1-+/3'A and SDF1-+/+ genotype survival. n=number of patients; RH=relative hazard and p=log likelihood p value based on the Cox proportional hazards model (Buchbinder, *AIDS* 8:1123, 1994; Lui *et al.*, *Science* 240:1333, 1988; Cox, and Oakes, *Analysis of Survival Data*, Chapman and Hall, London 1984, p. 36; Marubini and Valsecchi, *Analysing Survival Data for Clinical Trails and Observational Studies*, John Wiley, New York 1995 p. 160). SDF1 +/3'A survival was compared to SDF1-+/+ survival.

For every cohort, the SDF1-+/+ and SDF1-+/3'A individuals were indistinguishable in the pattern of progression to the three AIDS endpoints (Fig. 1A-I). There was however, a marked delay in AIDS onset among SDF1-3'A/3'A homozygotes with each AIDS endpoint. The difference was statistically significant in pooled (all) cohorts for AIDS-1993 and highly significant for AIDS-87 and death. Statistically significant relative hazards for the combined cohort (all races) analyses were notably low; namely 0.61 for AIDS-1993 (p=0.058), 0.34 for AIDS-1987 (p=0.002) and 0.23 for death (p=0.001). These values indicate that SDF1-+/+ and +/3'A individuals progress to AIDS 2-4 times more rapidly than SDF1-3'A/3'A homozygotes. Similar results were seen when Caucasian cohort members or different cohorts were analyzed separately (Table 1). SFCC, a cohort of homosexual men with a preponderance of long term survivors, had no deaths at all among SDF1-3'A/3'A homozygotes making estimates of relative hazard statistically less precise.

The extent of observed protection from AIDS progression associated with the SDF1-3'A/3'A genotype follows a gradation in combined and homosexual cohorts across increasingly severe AIDS endpoints (Fig. 1A-I). For example the RH value for the combined Caucasian cohort sample was 0.65 for AIDS-1993, 0.36 for AIDS-1987, and 0.24 for death (lower values indicate increased protection, Table 1). The tendency to display increased protection in later stages of HIV-1 infection was also seen in MACS and SFCC cohorts indicating that the protection increases as some patients progress to an AIDS-1993 definition (including drop of CD4 T lymphocytes to <200 cells/mm^3) the earliest step for most infected patients. This gradation was extended when time to CD4<200 cells mm^3 (alone without AIDS disease or death) was used as an endpoint, since in this case SDF1-3'A/3'A protection is barely detectable and not statistically significant. Protective genotypes include: *CCR2-64I* protection (*CCR5-+/+*, *CCR2-+/64I* or *64I/64I*, and *SDF1-+/+*); *CCR5-Δ32* protection (*CCR5-+/Δ32*, *CCR2-+/+*, and *SDF1-+/+*); *SDF1-3'A/3'A* plus at least one protective *CCR* allele (*SDF1-3'A/3'A* plus either *CCR5-+/Δ32* or *CCR2-+/64I*). Further, protective genotypes at either *CCR5* or *CCR2* are referred to as "CCR protection". This gradation indicates that the SDF-1-3'A/3'A protection is cummulative over the course of HIV-1 infection, and is possibly related to interference with the appearance of T-cell tropic HIV-1 populations.

Example 2

Protective effects of SDF-1 alleles

The protective effects of SDF1-3'A/3'A homozygotes were also apparent from a defined disease category analysis of SDF1 allele and genotype frequencies for the cohorts (Fig. 2A-F). Figure 2A-F are bar graphs which define disease category analysis of SDF1-3'A allele (panels A-C) and genotype (panels D-F) frequencies for each cohort and combined cohorts for the three endpoints, AIDS- 1993 (panels A and D), AIDS-1987 (panels B and E), and death (panels cand F). Seroconverters who progressed to the designated outcomes before the cutoff time were compared to seroconverters plus seroprevalents who survived outcome-free for at least that long.

Imputed seroconversion dates for the seroprevalent subgroup for MHCS, HGDS, and ALIVE were provided by the cohort investigators (Hilgartner *et al.*, *Am. J. Pediatr. Hematol. Oncol.* 15:208, 1993; Goedert *et al.*, *N. Engl. J. Med.* 321:1141, 1989; Lederman, *et al.*, *J. Infect. Dis.* 172:228, 1995; Vlahov *et al.*, *NIDA Research Monograph Series* 103 (Public Health Service, Alcohol and Drug Abuse Administration, Washington, DC, 1991). For MACS, date of enrollment was used at the starting date. Cutoffs, in years, were chosen as the time approximately half of all seroconverters had progressed to the outcomes. Times for the cutoffs were: 1) AIDS-1993; 7.5 year; 2) AIDS-1987, 8.5 years; 3) Death, 9.5 years. Individuals with outcome dates within one year of the cutoffs were not used in the analysis to remove patients with less certain infection-to-AIDS intervals due to imprecise seroconversion time windows (Cox, *J.R. Stat Soc. B* 34:187, 1972). The number of individuals in each disease category is listed below the bar graph. Categories with p-values in a Fisher's exact test (for the null hypothesis of left: allele frequency difference; or right: no SDF1-3'A/3'A protection compared to SDF1-+/+ plus SDF1-+/3'A) of less than 0.05 are marked*, and those below 0.01 with **. Bars are for Caucasians, triangles indicate SDF1-3'A/3'A frequencies for all racial groups. RR denotes relative risk of rapid progression for unprotected (SDF1-+/+ or SDF1-+/3'A) patients as compared to SDF1-3'A/3'A patients, among Caucasians in the combined cohorts; relative risks are calculated as case/control odds ratios taking slow progressors as controls; *i.e.*, the risk for each category is the ration of the number of rapid progressors to the number of slow progressors. 95% confidence intervals are in parentheses. P value is Fisher's exact test.

This analysis allows the inclusion of seroprevalent patients (those whose seroconversion data is unknown because they were HIV-1 antibody positive at the time of enrollment) in the slow/non-progressor category. The allele frequency of SDF1-3'A was higher among slow/non-progressors in combined and individual cohort analyses for each AIDS outcome (with one exception, SFCC due to extremely low numbers of rapid progressors), and several of the comparisons were statistically

significant. A separate examination of the frequencies of heterozygotes (+/3'A genotype) and homozygotes (3'A/3'A genotype) indicated that most of enrichment of the 3'A allele was contributed by the homozygotes. The heterozygotes +/3'A genotype was more frequent in the slow group in 10 of 12 cases, but this was significant for only one comparison (MHCS for death, P=.02, Fisher's exact test), while SDF1-3'A/3'A was more frequent in the slow group in all cases (Fig. 2A-F). The frequency of SDF1-3'A/3'A homozygotes was consistently higher among slow/non-progressors relative to rapid progressors, in many cases statistically significant for the three AIDS endpoints (Fig. 2A-F). The relative risk for AIDS occurrence (the fraction SDF1-+/+ and SDF1-+/3'A unprotected patients among those who progress to AIDS rapidly to the SDF1-3'A/3'A fraction who reach AIDS in a comparable period) ranged from 3.0-9.1 for the three AIDS endpoints (Fig. 2A-F). Not one SDF1-3'A/3'A homozygote was found among 63 patients from MHCS and SFCC that progressed to AIDS (by any definition) within 7.5 years compared to 4-5% frequency of SDF1-3'A/3'A homozygotes in those who avoid AIDS for 9.5 years or longer. The results of both the survival (Fig. 1A-I, Table 1) and the defined disease category analyses (Fig. 2) reveal a strong recessive SDF1-3'A association with protection against the clinical consequences of HIV-1 infection.

Common alleles within the coding region for the chemokine and M-tropic HIV-1 co-receptor genes, CCR5 and CCR2, have been shown to delay the rate of progression to AIDS (Samson, *Nature* 382:722, 1996; Dean *et al., Science* 273:1856, 1996; Huang *et al., Nature Med.* 2:1240, 1996; Michael *et al., Nature Med.* 3:338, 1997; Zimmerman *et al., Mol. Med.* 3:23, 1997; Biti *et al., Nature Med.* 3:252, 1997; O'Brien *et al., Lancet* 349:1219, 1997; Theodorou *et al., Lancet* 349:1219, 1997; Smith *et al., Science*, 277:959, 1997). The mutant alleles CCR5-Δ32 and CCR2-64I are dominant, genetically independent, and equally protective. An estimated 25-30% of long-term survivors who remain AIDS-free for > 16 years can be attributed to a protective genotype for either CCR5-Δ32 or CCR2-64I. A survival analysis of the relative contributions of CCR5-Δ32, CCR2-64I, and SDF1-

3'A genotypes (Fig. 3A-I, Table 2) reaffirm the protective effects of CCR2, CCR5 and SDF1 mutant genotypes on progression to AIDS when the influence of the other protective loci are considered as confounding variables.

Example 3

Protective effect of genotype

Figure 3A-I are graphs of Kaplan-Meier survival curves for the four protective genotypes for SDF1, CCR2, and CCR5 versus wild-type [+/+] on progression to AIDS-1993, AIDS-1987, and death in: the MACS cohort (panels A-C) ; Caucasians in the ALIVE, MACS, MHCS, SFCC combined cohorts (panels D-F) ; and all ethnic groups in the four combined cohorts (panels G-I). The protective genotypes are: SDF1-3'A/3'A; CCR2-[+/64I], [64I/64I]; and CCR5-+/Δ32. The four curves represent the following genotypes; 1) blue-+/+ at SDF1, CCR2 and CCR5; 2) green-CCR2/5 protein: one or more CCR2/5 protective genotypes and SDF-+/+ ; 3) orange-SDF1: SDF1-3'A/3'A and CCR2/5-+/+; 4) pink-SDF1 and CCR2/5: SDF1-3'A/3'A and protection by one or more CCR2/5 protective genotype versus +/+. n=number of individuals; p=log likelihood p value; and RH=relative hazard based on the Cox proportional hazards model (Center for Disease Control, *Morb. Mort. Wkly. Rep.* 41 (18 December 1992)). x-indicates single events; • indicate patient censoring. Summary statistics for each cohort and the combined cohort analyses are presented in Table 2.

The proportion of patients who progress rapidly or delay AIDS onset as a consequence of *SDF1* was estimated by computing the attributable risk of *SDF1* genotypes in extremely rapid (<3.5 yrs) and long term survivor (>16 yrs) disease categories. The *SDF-1* gene contains four exons over a 5.6 Kb region of chromosome 10q11.1 (Tashiro *et al.*, *Science*, 261:600, 1993). Two alternatively spliced transcripts which specify SDF-1α and SDF-1β are made from the gene and the isomers differ by the foreshortening of four carboxy terminal amino acids in SDF-1α (Tashiro *et al.*, *Science*, 261:600, 1993; Ross, *Trends Genet.*, 12:171, 1996); Tsai *et*

al., *J. Neurosci.*, 17:1950, 1997; McGowan *et al.*, *J. Biol. Chem.*, 272:1331, 1997; Shaw and Kamen, *Cell* 46:659, 1986; Kube *et al.*, *Cytokine* 7:107, 1995). The two transcripts have completely different 3'UTRs, and the *SDF1-3'A* mutation is found in the SDF1β transcript within a sequence block that is conserved between mouse and human *SDF1* UTR sequences. 1400 bp of SDF1β UTR were sequences and 2 additional mutations were discovered on both *SDF1-3'A* and *SDF1-+* allele containing haplotypes (Shirozu *et al.*, *Science* 28:495, 1995. The fraction of *SDF1-+/+* and *SDF1-+/3'A* individuals whose rapid progression could be attributed to their *SDF1* genotypes as compared to *SDF1-3'A/3'A* was 64% for AIDS-1993 (<5 years), 70% for AIDS 1987 (<5 years) and 81% for death (<8 years). Conversely, only 0.3%, 0.8% and 1.6% of the long term survivors (≥ 16 years) for AIDS 1993, AIDS 1987, and death, respectively could attribute their longevity to a *SDF1-3'A/3'A* genotype. The high attributable risks of *SDF1-+/+* and *+/3'A* in rapid progressors as compared to that seen in *SDF1-3'A/3'A* long term survivors is a consequence of the overall low frequency because attributable risk considers both strength of effect and frequency of the protective factor in the population of the *SDF1-3'A/3'A* genotype. Approximately 95% of the rapid progressors have permissive genotypes (*SDF1-+/+* or *+/3'A*) while only 4.9% of the slow non-progressors have the protective genotype (*SDF1-3'A/3'A*).). The RH values for SDF1-3'A/3'A genotype protection tend to be lower than for CCR protection (Table 2) suggesting that the SDF1-3'A/3'A genotype exhibits a stronger effect than either of the CCR protective genotypes. In addition CCR and SDF1 protection may be additive in AIDS cohorts, because patients with both SDF1 and CCR protective genotypes avoid AIDS outcomes longer than patients with only a single gene protection ($p = 0.05$ for AIDS-1993; $p < 0.01$ for AIDS-1987 and for death; Kaplan Meier log likelihood test). For example, only one of the ten seroconverter patients who were genotypically SDF1-3'A/3'A plus either CCR2 or CCR5 protected have progressed to AIDS defining pathologies (AIDS-1987), while eight SDF1-3'A/3'A, CCR2-+/+; CCR5-+/+ patients did (Fig. 3A-I). Ten dual protected (SDF1-3'A/3'A plus CCR2 or CCR5 protective genotype) seroprevalent

patients ultimately succumbed to AIDS, but their time interval from HIV-1 infection to AIDS was unknown.

When protective genotypes for all three loci are combined as a single genetic category and compared to non-protective genotypes (CCR5-+/+; CCR2-+/+; SDF1-+/+ or +/3'A), the results (Table 2) show significant protection for combined and MACS cohorts (Caucasian or combined ethnic groups) with significant relative hazards ranging from 0.44 to 0.65, suggesting that individuals with no protective genotypes progress to AIDS endpoints 60% more rapidly than HIV-1 infected individuals carrying protective alleles at one or more of the 3 loci.

Example 4

Cumulative protective effect

The cumulative effects of the SDF1-3'A/3'A protective genotype in combination (or not) with CCR protective genotypes were assessed over six intervals after HIV-1 seroconversion (Fig. 4A-C). Figure 4A-C shows the frequencies of the protective SDF1-3'A/3'A genotype alone (black) or in combination with at least one CCR2/5 protective genotype (CCR5-+/Δ32, CCR2-+/64I, and CCR2-64I/64I, cross hatch) in six intervals of increasing survivorship from midpoint (seroconverters) or imputed (seroprevalents) seroconversion dates in Caucasians. Genotypic frequencies were determined separately for time to AIDS-1993 (panel A), AIDS-1987 (panel B), and to death (panel C) using seroconverters progressing to the outcome in less than 3.5 years, and including seroconverters and seroprevalents progressing to the outcomes in 3.5 < 7 years, < 10 years, 10 < 13 years, and 13 < 16 years, and ≥16 years. The number of individuals observed in each category is shown above the column. The average frequency of the protective genotype for Caucasians is shown as an arrow. There is a statistically significant trend (Mantel-Haenzel X2) toward enrichment of SDF1-3'A/3'A genotypes at increasing survival intervals for AIDS 1993 (p=.03) and AIDS 1987 (p=0.02), and for increasing composite SDF-3'A/3'A plus CCR5 for

genotypes in longer survival intervals for AIDS-1993 (p=0.005), AIDS 87 (p=0.01) and death (p=0.04).

The results reveal a significant increase of SDF1-3'A/3'A genotypes among patients who avoid AIDS for longer periods. There was a complete absence of a dual CCR/SDF1 composite mutant genotypes among patients that develop AIDS-1987 or death within the first 10 years after HIV-1 infection and, only a single individual with SDF1 plus CCR protective genotypes developed AIDS-1993 during this interval. Combined with the survival analyses (Fig. 3, Table 2) the data emphasize the protective effect of the SDF1-3'A/3'A genotype and suggest its effect is additive with the protection afforded by CCR2 and CCR5 variant alleles (Levin, *Acta Inter. Canc.* 9:531, 1953). Estimates of attributable risk were computed for Caucasians (Fig. 4), because of the differences in *CCR5-Δ32* and *SDF1-3'A* allele frequencies in Asian and African individuals (Dean *et al.*, *Science* 273:1856, 1996; Smith *et al.*, *Science*, 277:959, 1997).

The finding that SDF1-3'A/3'A homozygotes postpone AIDS onset raises several issues about the mechanism of viral restriction and AIDS pathogenesis. The SDF1-3'A mutation is located in the 3' untranslated region of the SDF1β gene transcript . A screen of 8 homozygous SDF1-3'A/3'A individuals for mutations in the four exons did not reveal any additional polymorphisms. Absence of intragenic variants is not surprising since the SDF1 gene is highly conserved among mammals with only one amino acid difference observed between human and mouse homologues. It is conceivable that the 3'UTR variant may influence mRNA transcript synthesis, persistence, transport, splice product abundance or response to transcription factors as have been reported for other 3'UTR systems (Mellors *et al.*, *Science* 272:1167, 1996; O'Brien *et al.*, *J. Am. Med. Assoc.* 276:105, 1996). The SDF1 gene specifies alternatively spliced transcripts SDF-1α and SDF-1β that differ by the loss in SDF-1α of four carboxy terminal amino acids and in their 3'UTR sequence. The development of specific reagents to assess these chemokine gene products and their

transcripts would permit the investigation of SDF1 genotype influence on SDF-1 protein availability and function.

If the SDF1 mutation would affect the availability, persistence, or CXCR4 receptor avidity of the chemokine, then the avoidance of AIDS symptoms could be interpreted as a consequence of limiting the emergence and spread of the pathogenic T-tropic HIV-1 (Fauci, *Nature* 384:529, 1996; Weiss, *Science* 272:1885, 1996; Feng *et al.*, *Science* 272:872, 1996; Godfrey-Faussett *et al.*, *Nature* 368:183, 1994; Williams, CRC press. Ann Arbor, (1992) pp.363; DeCock, *Science* 249:793, 1990). Under this hypothesis, the robust but recessive character of SDF1-3'A/3'A protection makes a prediction that the SDF1-3'A specific protein is more effective than the SDF-1+ product in restricting late stage T-tropic viral load, a strong correlate and prognostic indicator of pathogenesis. The SDF1 alteration may impede AIDS onset simply by reducing availability of requisite CXCR4 co-receptors to HIV-1 by up-regulation of SDF1, by increasing SDF-1 transcript stability, or by other interactions. Such mechanisms would also account for the gradation in survival outcomes whereby SDF1-3'A/3'A effects are more pronounced in late stage outcomes (AIDS-1987 and death) than are evident with earlier stages of HIV-1 infection (Fig. 1A-I).

SDF1 protection is recessive, making the prevalence of protected individuals below 5%. The combination of CCR2 or CCR5 plus SDF1 protection is even more powerful (Fig. 4), but also very rare (< 2% in non African ethnic groups and < 0.1% in Africans). These low frequencies indicate that the restriction gene variants would exert only a slight effect on the progress of the AIDS epidemic, particularly in Africa where AIDS is spreading at a high rate and where protective alleles at CCR5 and SDF1 are even less frequent. Nonetheless, the finding that in the groups examined SDF1-3'A/3'A genotypes protect almost completely from rapid progression and that 70-80% of the AIDS patients who progress to AIDS rapidly do so because they lack SDF1 protection raise the possibility of drug or gene therapy that would mimic SDF1 or composite SDF-1 plus CCR protection to HIV-1 exposed and infected individuals.

Table 1 shows a survival analysis for progression to 3 AIDS endpoints among HIV-1 infected seroconverters for SDF1-3'A/3'A versus SDF1-+/+ or SDF-+/3'A genotypes as in Figure 1A-I. Seroconverters for the ALIVE, MACS, SFCC, MHCS, and combined cohorts including only Caucasians, and for the combined cohorts with all ethnic groups included were analyzed using the Cox proportional hazards model. The HGDS cohort was excluded since all the participants were HIV-1 infected prior to study entry. ALIVE is not analyzed as a separate cohort because the combination of the recency of cohort (1988) plus the low allele frequency among African Americans which comprise 94% of the cohort resulted in too few SDF1-3'A/3'A AIDS outcomes to be statistically robust. A log likelihood test (1df) (LL), p value, and relative hazard (RH) were calculated for each variable in the analysis of AIDS outcomes. Time to AIDS-1993, AIDS-1987 and death were calculated from the midpoint of the last HIV-1 negative test date and the first HIV-1 positive test date. Seroconverters with an interval greater than 3 years between last negative and first positive were excluded from the analysis. Analyses were adjusted for age, where age is a categorical variable with three categories: < 30, 30-40, or > 40 yr. old. A Bonferonni correction for multiple tests indicates levels of significance *$p \leq 0.05$; ** $p \leq 0.01$.

Table 1

	All Ethnic Groups					Caucasians				
Cohort	n	events	RH	LL	p-value	n	events	RH	LL	p-value
AIDS 1993										
MACS	387	230	0.59	2.27	0.13	329	201	0.59	2.15	0.14
MHCS	156	110	0.38	2.46	0.12	140	97	0.39	2.34	0.13
SFCC	171	114	0.69	0.44	0.51	157	102	0.67	0.51	0.47
All	843	504	0.58	4.22	0.04	635	401	0.61	3.23	0.07
AIDS 1987										
MACS	390	198	0.22	11.19	0.0008***	332	175	0.22	10.95	0.0009**
MHCS	156	74	0.58	0.69	0.41	140	69	0.59	0.63	0.43
SFCC	171	74	0.31	2.04	0.15	157	70	0.30	2.14	0.14
All	856	386	0.34	9.74	0.002**	638	314	0.36	8.76	0.003**
Death										
MACS	390	153	0.10	12.63	0.0004**	332	138	0.10	12.82	0.0003**
MHCS	156	58	0.84	0.06	0.81	140	55	0.85	0.06	0.81
SFCC	171	51	0.0	4.78	0.03	157	49	0.0	5.00	0.03
All	856	293	0.23	10.91	0.001**	638	242	0.24	10.12	0.002**

Relative hazards for SDF1-3'A/3'A versus SDF1-+/+ genotypes.

Table 2 shows a survival analysis of protection from progression to AIDS outcomes by SDF1-3'A/3'A variant, CCR5 or CCR2 protective polymorphisms, and a second analysis of any protection by any variant at CCR5, CCR2 and SDF1. The analyses using the Cox proportional hazards model were performed as in Table 1. Protective genotypes at SDF1 were considered to be 3'A/3'A vs. 3'A/+ or +/+; CCR2; 641/641 or 641/+ vs. +/+; CCR5-+/Δ32 or Δ32/Δ32 vs. +/+. The SDF1 genotypes were analyzed three ways: 1) SDF1-3'A/3'A versus SDF1-+/+ or SDF1-+/3'A controlling for the protective genotypes of CCR2 and CCR5. 2) CCR2-641/641 or CCR2-+/641 or CCR5-+/Δ32 versus CCR5-+/+ and CCR2-+/+ (normal at tow loci) controlling for the protective genotype of SDF1. 3) SDF1-3'A/3'A and/or one or more protective CCR2/5 genotypes versus +/+ at all three loci. Individual and combined cohorts consider Caucasians except for "all races" where all ethnic groups are analyzed. p-values for the log likelihood tests (LL) employ a Bonferroni correction for multiple tests performed in each of the three analyses: *=P$\leq$ 0.01 and **=P$\leq$0.001, and P $\leq$ .0001***. A log likelihood calculation for X2 was performed because of the small numbers of patients and few failures in SDF1-3'A/3'A individuals.†[+/+] for SDF1 includes SDF1-+/+ and SDF1-+/3'A.

Table 2

SDF1 and CCR protective effects

			SDF1-3'A/3'A vs. [+/+] for SDF1, CCR2, CCR5			CCR2-+/641 or CCR5-+/Δ32 vs. [+/+] for SDFI, CCR2 and CCR5			Any vs. no protective genotype(s)		
cohort	n	events	RH	LL	P	RH	LL	P	RH	LL	P
AIDS 1993											
MACS	313	195	0.63	1.7	0.20	0.75	3.4	0.06	0.76	3.4	0.07
MHCS	124	84	0.34	3.2	0.08	0.45	10.4	0.001**	0.44	11.7	0.0006***
SFCC	156	102	0.69	0.5	0.50	0.57	7.4	0.006	0.58	7.4	0.007
all Cauc.	601	382	0.64	2.7	0.10	0.61	21.5	0.000003***	0.61	21.4	0.000004**
all races	798	481	0.60	3.6	0.06	0.64	21.7	0.000003***	0.65	22.3	0.000002**
AIDS 1987											
MACS	316	172	0.23	10.7	0.001***	0.80	1.81	0.18	0.71	4.6	0.03
MHCS	124	60	0.60	0.6	0.45	0.49	5.5	0.02	0.52	5.2	0.02
SFCC	156	70	0.32	1.9	0.17	0.57	5.1	0.02	0.56	5.6	0.02
all Cauc.	604	302	0.37	8.3	0.004*	0.64	13.7	0.0002**	0.61	17.1	0.00004
all races	804	370	0.35	9.4	0.002*	0.66	13.7	0.0002**	0.63	17.3	0.00003***
DEATH											
MACS	316	137	0.10	13.0	0.0003**	0.63	6.4	0.01	0.55	10.9	0.001**
MHCS	124	47	0.96	0.004	0.95	0.37	7.2	0.007	0.44	0.37	0.02
SFCC	156	49	NA	5.01	0.03	0.65	2.1	0.14	0.58	0.65	0.07
all Cauc.	604	233	0.24	10.0	0.002*	0.55	17.8	0.00003***	0.52	0.55	0.000003**
all races	804	281	0.23	10.9	0.001**	0.59	16.1	0.00006***	0.56	0.59	0.000006**

It will be apparent to those skilled in the art that various modifications and variations can be made to the compounds and processes of this invention. Thus, it is intended that the present invention cover such modifications and variations, provided they come within the scope of the appended claims and their equivalents. Accordingly, the invention is limited only by the following claims.

What is claimed is:

1. An isolated polynucleotide encoding a stromal cell derived factor-1 (SDF-1) variant having a nucleotide sequence set forth in SEQ ID NO:1.

2. A method for determining the prognosis of a subject exposed to HIV-1 comprising:
   a) determining the presence of a SDF-1 variant nucleic acid in cells of said subject; and
   b) correlating the presence of the variant on both alleles with prognosis of said subject.

3. The method of claim 2, wherein said cells are peripheral blood leukocytes (PBLs).

4. The method of claim 2, further comprising determining the presence of a CCR2 mutation .

5. The method of claim 2, further comprising determining the presence of a CCR5 mutation .

6. The method of claim 4, further comprising determining the presence of a CCR5 mutation .

7. A method of determining susceptibility of a subject to HIV infection comprising determining the SDF-1 allelic profile of a subject comprising the isolating the SDF-1 nucleic acid sequence and determining the presence or absence of a mutation in SDF-1 nucleic acid.

8. The method of claim 7, wherein the nucleic acid is amplified after isolating.

9. The method of claim 7, wherein the SDF-1 mutation is a G to A transition mutation at nucleotide 801.

10. The method of claim 7, further comprising determining the CCR2 allelic profile of the subject.

11. The method of claim 7, further comprising determining the CCR5 allelic profile of the subject.

12. The method of claim 10, further comprising determining the CCR5 allelic profile of the subject.

13. A method of inhibiting membrane fusion between HIV and a target cell that expresses CXCR4 or between an HIV-infected cell and a CD4 positive uninfected cell that expresses CXCR4 comprising contacting the target or CD4/CXCR4 positive cell with a CXCR4 down-regulating effective amount of a SDF-1 variant, thereby inhibiting membrane fusion.

14. The method of claim 13, wherein the SDF-1 variant has the nucleotide sequence as set forth in SEQ ID NO:1.

15. The method of claim 13, wherein the contacting is by *in vivo* administration to a subject.

16. The method of claim 13, wherein the contacting is by *ex vivo* administration to a cell.

17. A method of treating a subject having or at risk of having an HIV infection or disorder, comprising administering to the subject, a therapeutically effective amount of an SDF-1 variant.

18. The method of claim 17, wherein the SDF-1 variant is the nucleotide sequence as set forth in SEQ ID NO:1.

19. The method of claim 17, wherein the SDF-1 variant is administered to a subject suffering from AIDS or ARC.

20. A method of treating a subject having a disorder associated with expression of CXCR4 comprising administering to the subject, an SDF-1 variant that suppresses CXCR4.

21. The method of claim 20, wherein the SDF-1 variant is introduced into the cell using a carrier.

22. The method of claim 21, wherein the carrier is a vector.

23. The method of claim 20, wherein the administering is *ex vivo*.

24. The method of claim 20, wherein the administering is *in vivo*.

25. A kit for determining the SDF-1 allelic profile of an individual, comprising amplification primers or hybridization probes which detect a transition mutation of G to A at nucleotide 801.

FIGURE 1

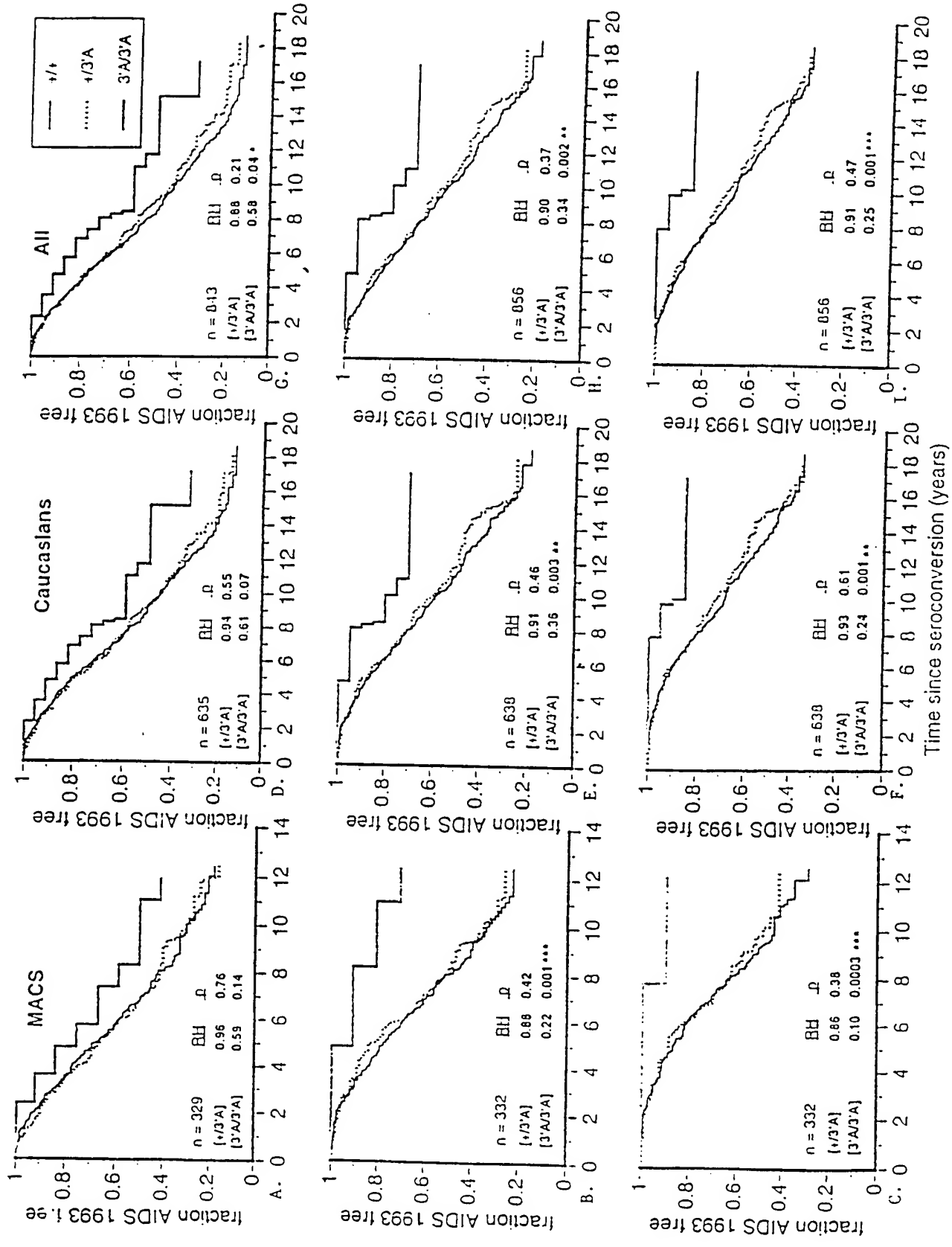
+/+
+/3'A
3'A/3'A
All
MACS
Caucaslans
Fraction AIDS 1993 free
Time since seroconversion (years)
A.
B.
C.
D.
E.
F.
G.
H.
I.

FIGURE 2

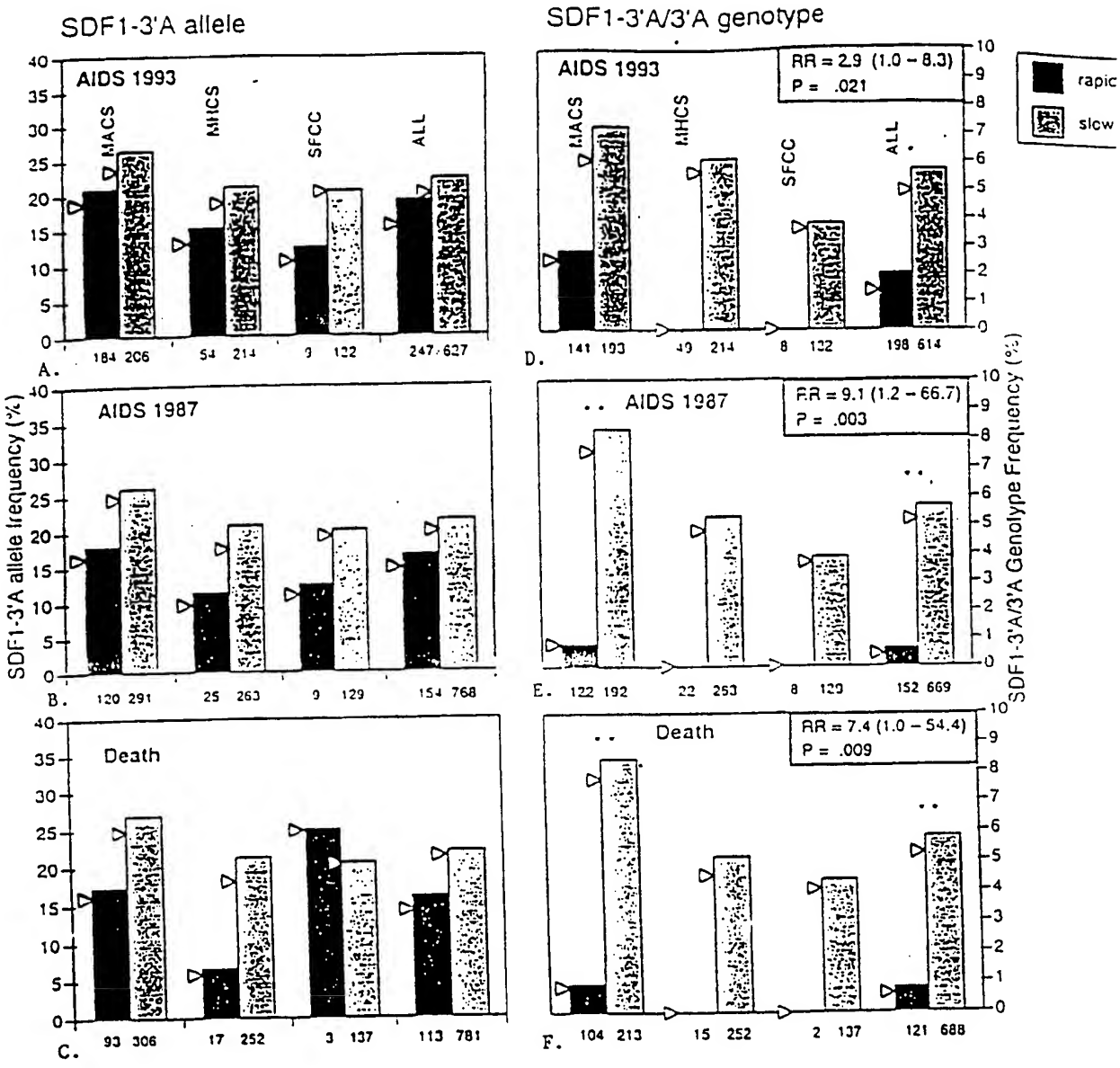
SDF1-3'A allele
SDF1-3'A/3'A genotype
AIDS 1993
MACS
MHCS
SFCC
ALL
RR = 2.9 (1.0 – 8.3)
P = .021
rapic
slow
A. 184 206 54 214 9 122 247 627
D. 141 193 49 214 8 132 198 614
AIDS 1987
RR = 9.1 (1.2 – 66.7)
P = .003
B. 120 291 25 263 9 129 154 768
E. 122 192 22 253 8 129 152 669
Death
RR = 7.4 (1.0 – 54.4)
P = .009
C. 93 306 17 252 3 137 113 781
F. 104 213 15 252 2 137 121 688
SDF1-3'A allele frequency (%)
SDF1-3'A/3'A Genotype Frequency (%)

FIGURE 3

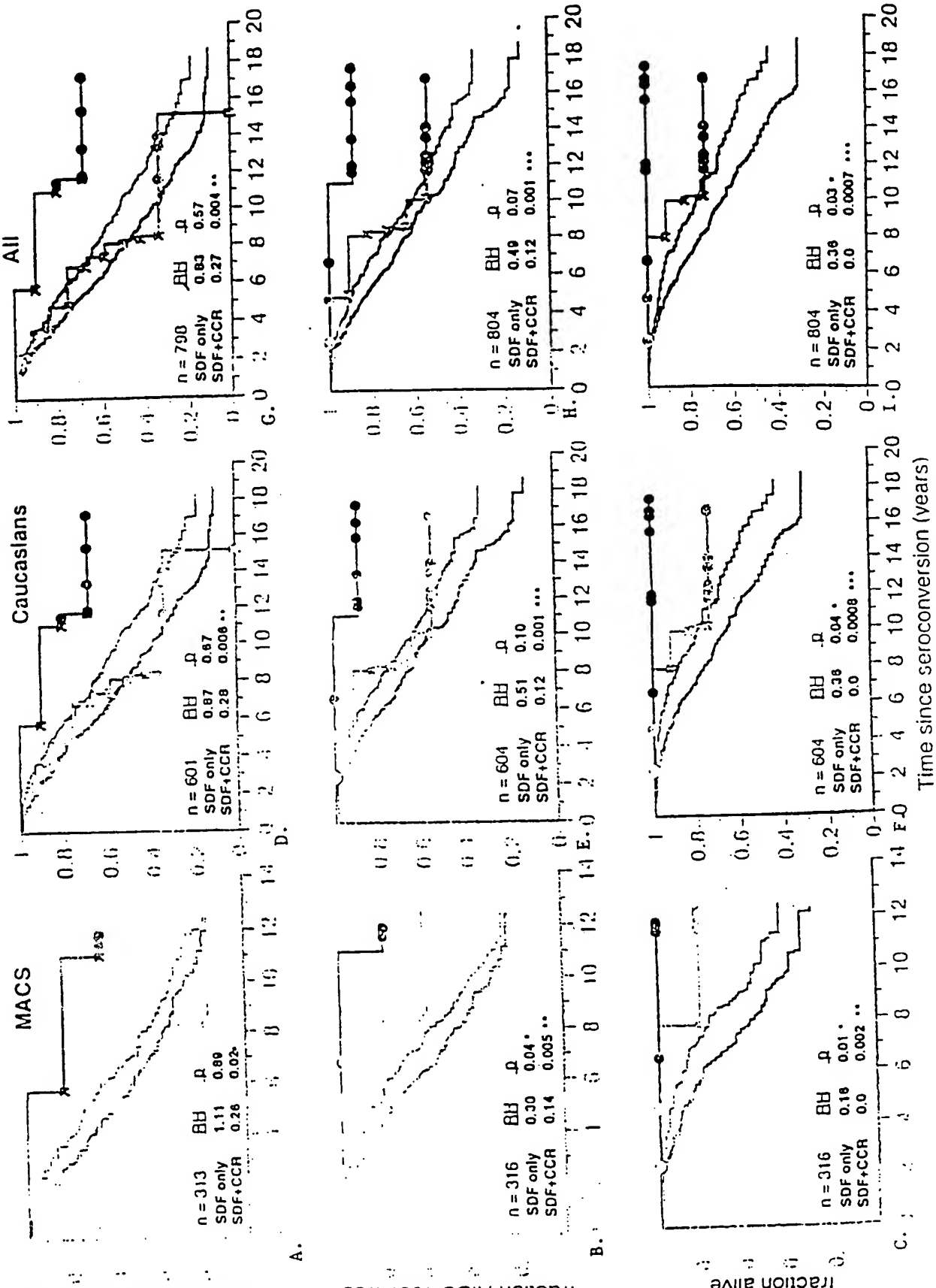

FIGURE 4

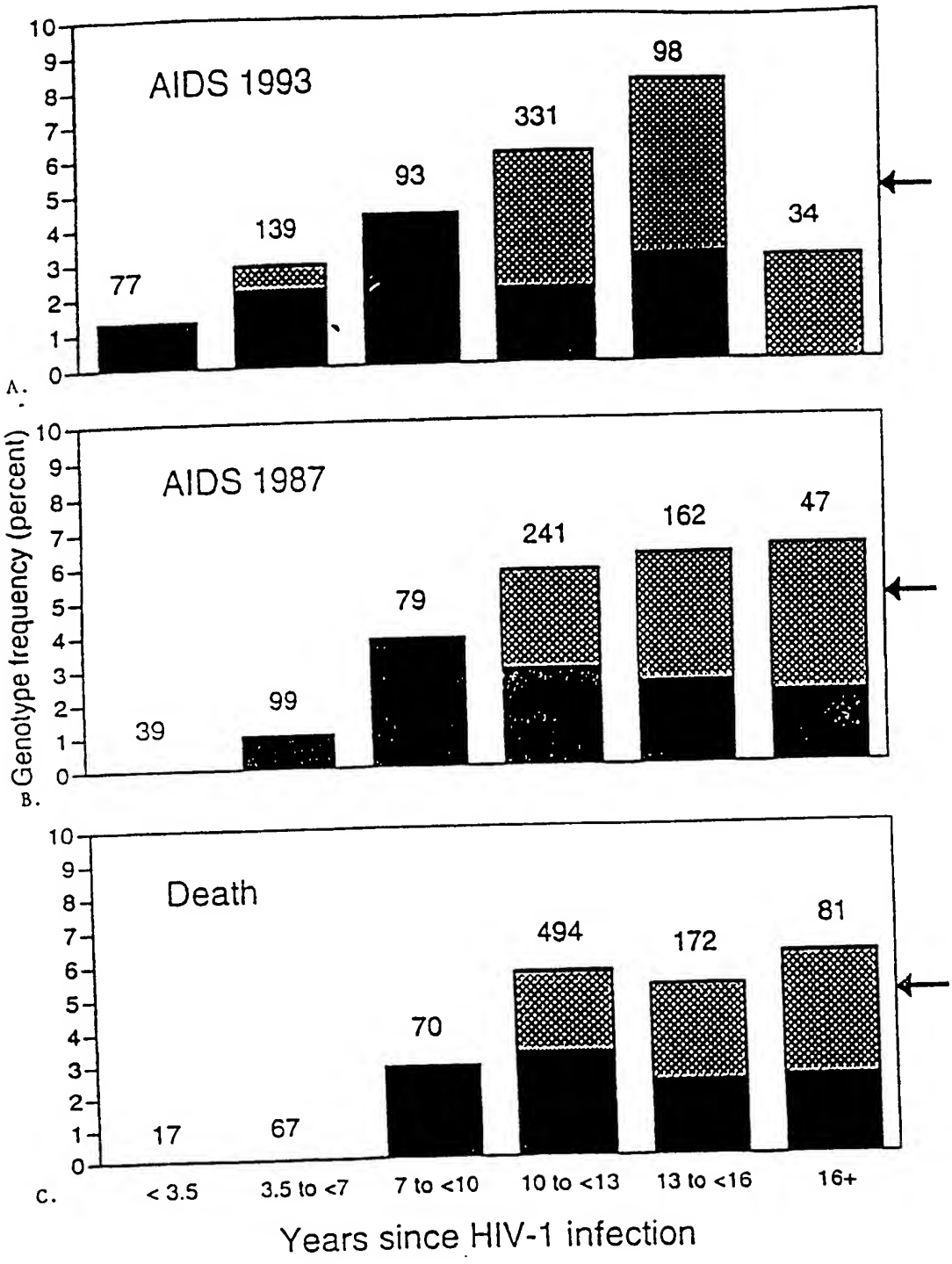
AIDS 1993
77
139
93
331
98
34
A.
AIDS 1987
39
99
79
241
162
47
B.
Genotype frequency (percent)
Death
17
67
70
494
172
81
C.
< 3.5
3.5 to <7
7 to <10
10 to <13
13 to <16
16+
Years since HIV-1 infection

FIGURE 5 A

SDF-1 B - 3'A Mutation: (G801A) (SEQ ID NO:1)

```
   1 tctccgtcag ccgcattgcc cgctcggcgt ccggcccccg acccgtgctc gtccgcccgc
  61 ccgcccgccc gcccgcgcca tgaacgccaa ggtcgtggtc gtgctggtcc tcgtgctgac
 121 cgcgctctgc ctcagcgacg ggaagcccgt cagcctgagc tacagatgcc catgccgatt
 181 cttcgaaagc catgttgcca gagccaacgt caagcatctc aaaattctca acactccaaa
 241 ctgtgccctt cagattgtag cccggctgaa gaacaacaac agacaagtgt gcattgaccc
 301 gaagctaaag tggattcagg agtacctgga gaaagcttta aacaagaggt tcaagatgtg
 361 agagggtcag acgcctgagg aacccttaca gtaggagccc agctctgaaa ccagtgttag
 421 ggaagggcct gccacagcct cccctgccag ggcagggccc caggcattgc caagggcttt
 481 gttttgcaca ctttgccata ttttcaccat ttgattatgt agcaaaatac atgacattta
 541 ttttttcattt agtttgatta ttcagtgtca ctggcgacac gtagcagctt agactaaggc
 601 cattattgta cttgccttat tagagtgtct ttccacggag ccactcctct gactcagggc
 661 tcctgggttt tgtattctct gagctgtgca ggtggggaga ctgggctgag ggagcctggc
 721 cccatggtca gccctagggt ggagagccac caagagggac gcctgggggt gccaggacca
 781 gtcaacctgg gcaaagccta gtgaaggctt ctctctgtgg gatgggatgg tggagggcca
 841 catgggaggc tcacccccct ctccatccac atgggagccA ggtctgcctc ttctgggagg
 901 gcagcagggc taccctgagc tgaggcagca gtgtgaggcc agggcagagt gagacccagc
 961 cctcatcccg agcacctcca catcctccac gttctgctca tcattctctg tctcatccat
1021 catcatgtgt gtccacgact gtctccatgg ccccgcaaaa ggactctcag gaccaaagct
1081 ttcatgtaaa ctgtgcacca agcaggaaat gaaaatgtct tgtgttacct gaaaacactg
1141 tgcacatctg tgtcttgttt ggaatattgt ccattgtcca atcctatgtt tttgttcaaa
1201 gccagcgtcc tcctctgtga ccaatgtctt gatgcatgca ctgttccccc tgtgcagccg
1261 ctgagcgagg agatgctcct tgggcccttt gagtgcagtc ctgatcagag ccgtggtcct
1321 ttggggtgaa ctaccttggt tccccccactg atcacaaaaa catggtgggt ccatgggcag
1381 agcccaaggg aattcggtgt gcaccagggt tgaccccaga ggattgctgc cccatcagtg
1441 ctcccctcaca tgtcagtacc ttcaaactag ggccaagccc agcactgctt gaggaaaaca
1501 agcattcaca acttgttttt ggtttttaaa acccagtcca caaaataacc aatcctggac
1561 atgaagattc tttcccaatt cacatctaac ctcatcttct tcaccatttg gcaatgccat
1621 catctcctgc cttcctcctg ggccctctct gctctgcgtg tcacctgtgc ttcgggccct
1681 tcccacagga catttctcta agagaacaat gtgctatgtg aagagtaagt caacctgcct
1741 gacatttgga gtgttccccc tccactgagg gcagtcgata gagctgtatt aagccactta
1801 aaatgttcac ttttgacaaa ggcaagcact tgtgggtttt tgttttgttt ttcattcagt
1861 cttacgaata cttttgccct ttgattaaag actccagtta aaaaaaattt taatgaagaa
1921 agtggaaaac aaggaagtca aagcaaggaa actatgtaac atgtaggaag taggaagtaa
1981 attatagtga tgtaatcttg aattgtaact gttcttgaat ttaataatct gtagggtaat
2041 tagtaacatg tgttaagtat tttcataagt atttcaaatt ggagcttcat ggcagaaggc
2101 aaacccatca acaaaaattg tcccttaaac aaaaattaaa atcctcaatc cagctatgtt
2161 atattgaaaa aatagagcct gagggatctt tactagttat aaagatacag aactctttca
2221 aaacctttg aaattaacct ctcactatac cagtataatt gagttttcag tggggcagtc
2281 attatccagg taatccaaga tattttaaaa tctgtcacgt agaacttgga tgtacctgcc
2341 cccaatccat gaaccaagac cattgaattc ttggttgagg aaacaaacat gaccctaaat
2401 cttgactaca gtcaggaaag gaatcatttc tatttctcct ccatgggaga aaatagataa
2461 gagtagaaac tgcagggaaa attatttgca taacaattcc tctactaaca atcagctcct
2521 tcctggagac tgcccagcta aagcaatatg catttaaata cagtcttcca tttgcaaggg
2581 aaaagtctct tgtaatccga atctcttttt gctttcgaac tgctagtcaa gtgcgtccac
2641 gagctgttta ctagggatcc ctcatctgtc cctccgggac ctggtgctgc ctctacctga
2701 cactcccttg ggctcccctgt aacctcttca gaggccctcg ctgccagctc tgtatcagga
2761 cccagaggaa ggggccagag gctcgttgac tggctgtgtg ttgggattga gtctgtgcca
2821 cgtgttttgtg ctgtggtgtg tccccctctg tccaggcact gagataccag cgaggaggct
2881 ccagagggca ctctgcttgt tattagagat tacctcctga gaaaaaagct tccgcttgga
2941 gcagaggggc tgaatagcag aaggttgcac ctccccccaac cttagatgtt ctaagtcttt
```

FIGURE 5 B

```
3001 ccattggatc tcattggacc cttccatggt gtgatcgtct gactggtgtt atcaccgtgg
3061 gctcccctgac tgggagttga tcgccctttcc caggtgctac acccctttcc agctggatga
3121 gaatttgagt gctctgatcc ctctacagag cttcccctgac tcattctgaa ggagccccat
3181 tcctgggaaa tattcccctag aaacttccaa atcccctaag cagaccactg ataaaaccat
3241 gtagaaaatt tgttattttg caacctcgct ggactctcag tctctgagca gtgaatgatt
3301 cagtgttaaa tgtgatgaat actgtatttt gtattgtttc aagtgcatct cccagataat
3361 gtgaaaatgg tccaggagaa ggccaattcc tatacgcagc gtgctttaaa aaataaataa
3421 gaaacaactc tttgagaaac aacaatttct actttgaagt cataccaatg aaaaaatgta
3481 tatgcactta taatttttcct aataaagttc tgtactcaaa tgta
```

(Circled nucleotide) G801A (3'A) mutation associated with delayed progression.

Mistakes in the published sequence L36033 are corrected and underlined.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/22578

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C12Q 1/68; A61K 38/19; C07H 21/02, 21/04.
US CL :435/6; 514/2, 8, 21; 536.23.5, 24.3, 24.31.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/6; 514/2, 8, 21; 536.23.5, 24.3, 24.31.

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, DIALOG

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X -- Y	UEDA et al. Chemically synthesized SDF-1alpha analogue, N33A, is a potent chemotactic agent for CXCR4/Fusin/LESTR- expressing human leukocytes. The Journal of Biological Chemistry. 03 October 1997, Vol. 272, No. 40, pages 24966-24970, see entire document, especially paragraph bridging pages 24969-24970.	13, 16 -------- 1, 14-15, 25
X, P ---- Y, P	LUSSO, P. A chemokine trap for HIV co-receptors. Nature Medicine. October 1997, Vol. 3, No. 10, pages 1074-1075, see entire document.	13, 16, ---------- 1, 14-15, 25
X, P ---- Y, P	WINKLER et al. Genetic restriction of AIDS pathogenesis by an SDF-1 chemokine gene variant. Science. 16 January 1998, Vol. 279, pages 389-393, see entire document.	1, 25 ------- 13-16

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search
18 DECEMBER 1998

Date of mailing of the international search report
02 FEB 1999

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703) 305-3230

Authorized officer
ROBERT D. BUDENS
Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATI L SEARCH REPORT

International application No.
CT/US98/22578

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X, P ---- Y, P	MICHAEL, N.L. Protective effect of an SDF-1 variant in HIV disease. Journal of the American Medical Association. 08 April 1998, Vol. 279, No. 14, page 1140, see entire document.	1, 25 ------ 13-16
Y	BLEUL et al. The lymphocyte chemoattractant SDF-1 is a ligand for LESTR/fusin and blocks HIV-1 entry. Nature. 29 August 1996, Vol. 382, pages 829-833, see entire document.	1, 13-16, 25
Y	DORANZ et al. A small-molecule inhibitor directed against the chemokine receptor CXCR4 prevents its use as an HIV-1 coreceptor. Journal of Experimental Medicine. 20 October 1997, Vol. 186, No. 8, pages 1395-1400, see entire document.	1, 13-16, 25
Y, P	CRUMP et al. Solution structure and basis for functional activity of stromal cell-derived factor-1; dissociation of CXCR4 activation from binding and inhibition of HIV-1. The EMBO Journal. 1997, Vol. 16, No. 23, pages 6996-7007, see entire document.	1, 13-16, 25
A	DAAR et al. High concentrations of recombinant soluble CD4 are required to neutralize primary human immunodeficiency virus type 1 isolates. Proceedings of the National Academy of Sciences, USA. September 1990, Vol. 87, pages 6574-6578, see entire document.	1, 13-16, 25

International application No.
PCT/US98/22578

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s) 1 and 13-16, drawn to a polynucleotide and method of use.
Group II, claim(s) 2-6, drawn to methods of determining prognosis of HIV infection.
Group III, claim(s) 7-12, drawn to methods of determining susceptibility to HIV infection.
Group IV, claim(s)17-19, drawn to methods of treating HIV infection.
Group V, claim(s) 20-24, drawn to methods of treating CXCR4 disorders.
Group VI, claim(s) 25, drawn to a test kit for SDF-1 variants.

The inventions listed as Groups I-V do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: the methods of Groups I-V differ in their industrial applicability and in their method steps and reagents and are not so linked by a special technical feature within the meaning of PCT Rule 13.2 so as to form a single general inventive concept.

The inventions listed as Groups I and VI do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: the products of Groups I and VI differ in their physical and chemical properties. The product of Group I is directed to a nucleotide encoding an SDF-1 variant while the product of Group VI only encompasses probes which could be used to amplify or hybridize to polynucleotides encoding SDF-1. Thus the inventions of Groups I and VI are not so linked by a special technical feature within the meaning of PCT Rule 13.2 so as to form a single general inventive concept.

FIGURE 1

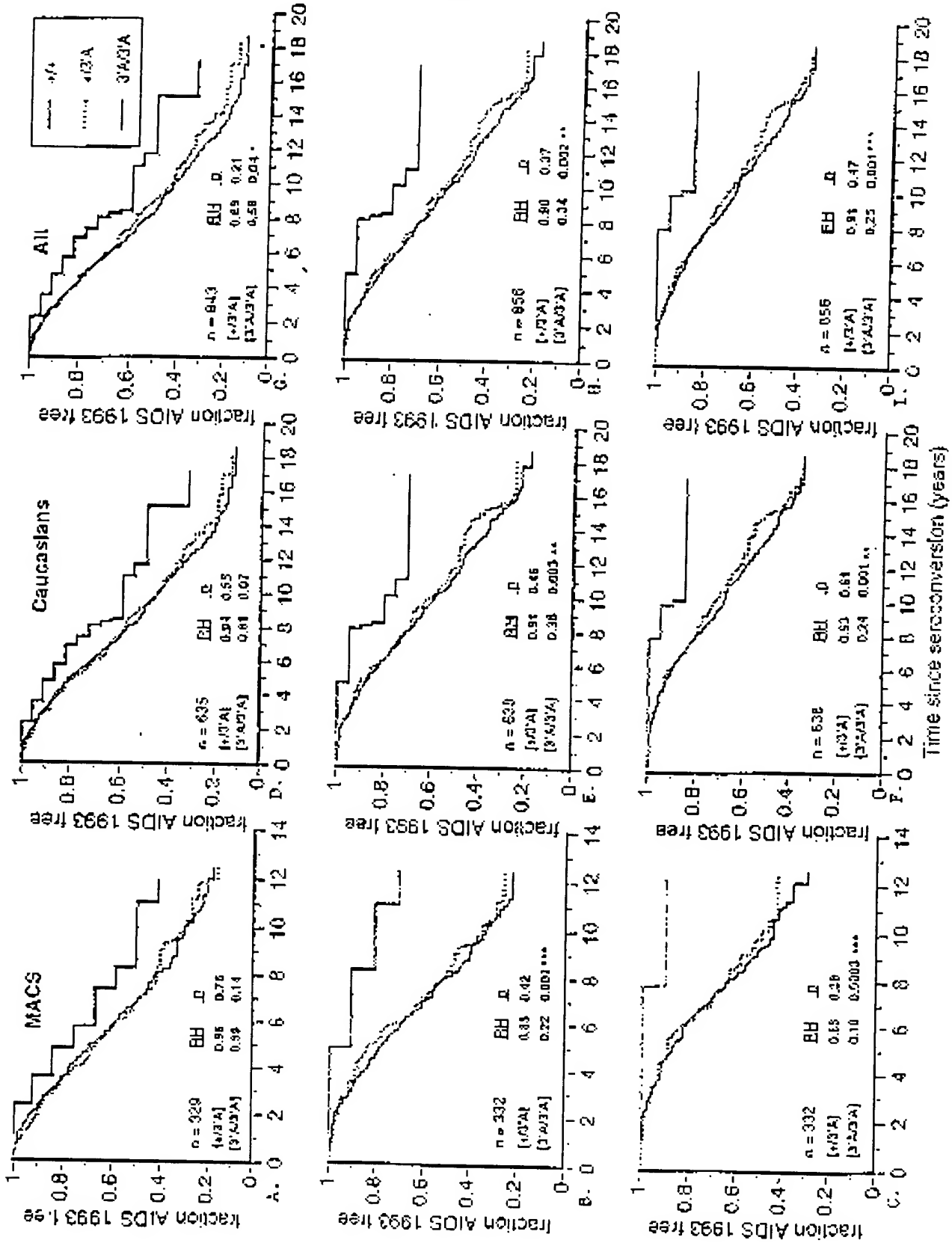
All
MACS
Caucasians
Fraction AIDS 1993 tree
Time since seroconversion (years)

FIGURE 2

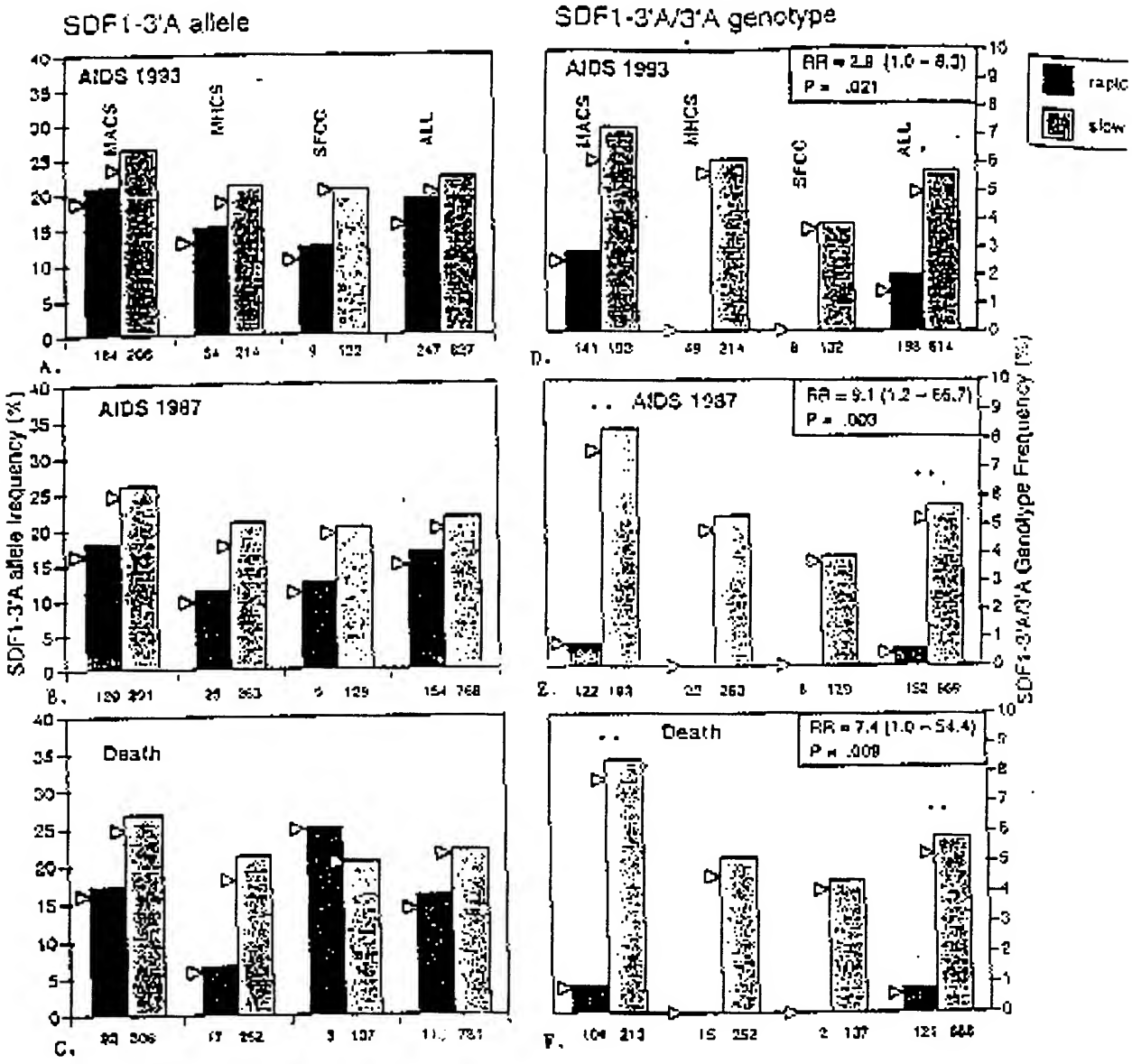
SDF1-3'A allele
SDF1-3'A/3'A genotype
AIDS 1993
AIDS 1987
Death
MACS
MHCS
SFCC
ALL
RR = 2.8 (1.0 – 8.3)
P = .021
RR = 9.1 (1.2 – 66.7)
P = .003
RR = 7.4 (1.0 – 54.4)
P = .009
rapid
slow
SDF1-3'A allele frequency (%)
SDF1-3'A/3'A Genotype Frequency (%)
A.
B.
C.
D.
E.
F.

FIGURE 3

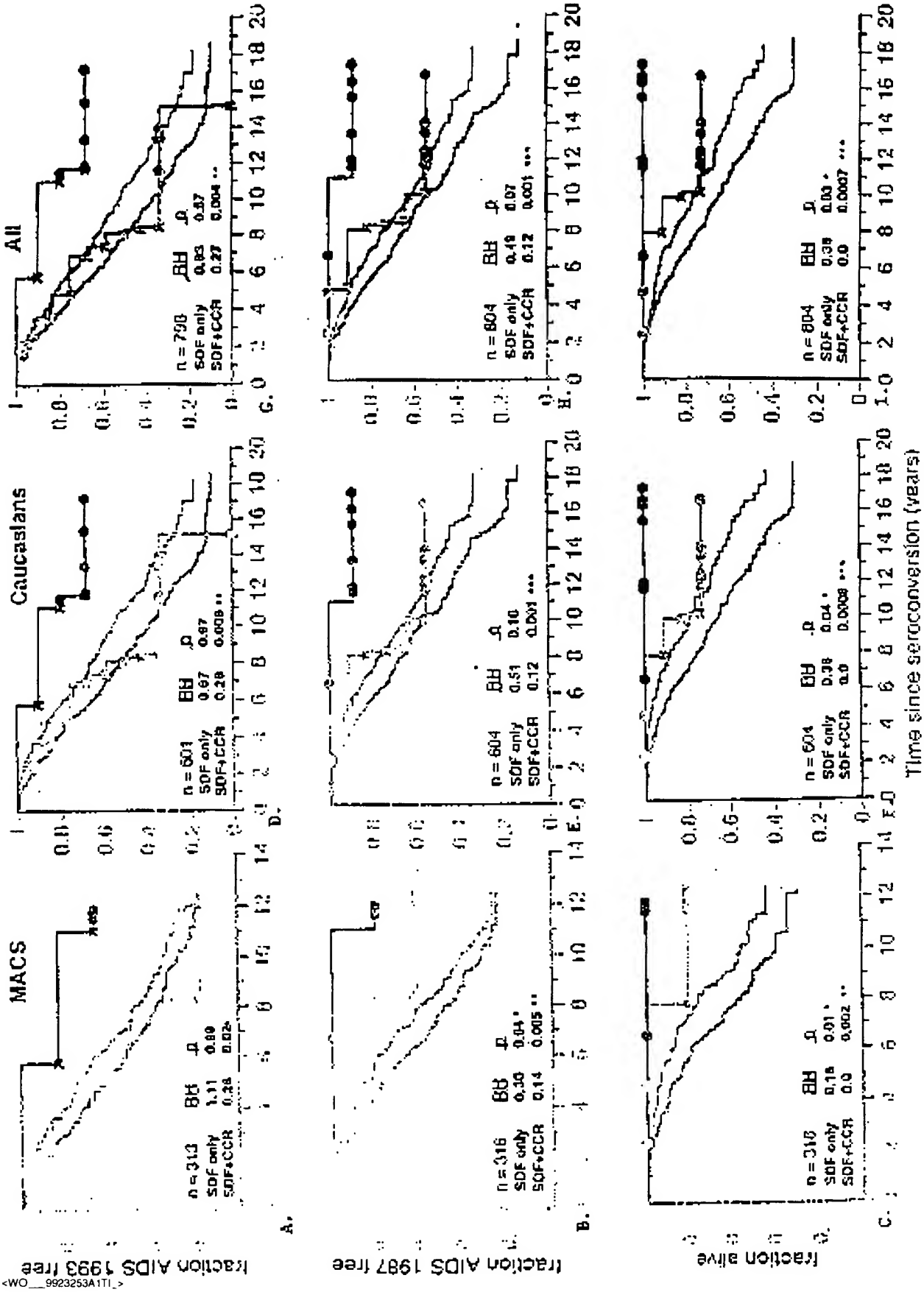

FIGURE 4

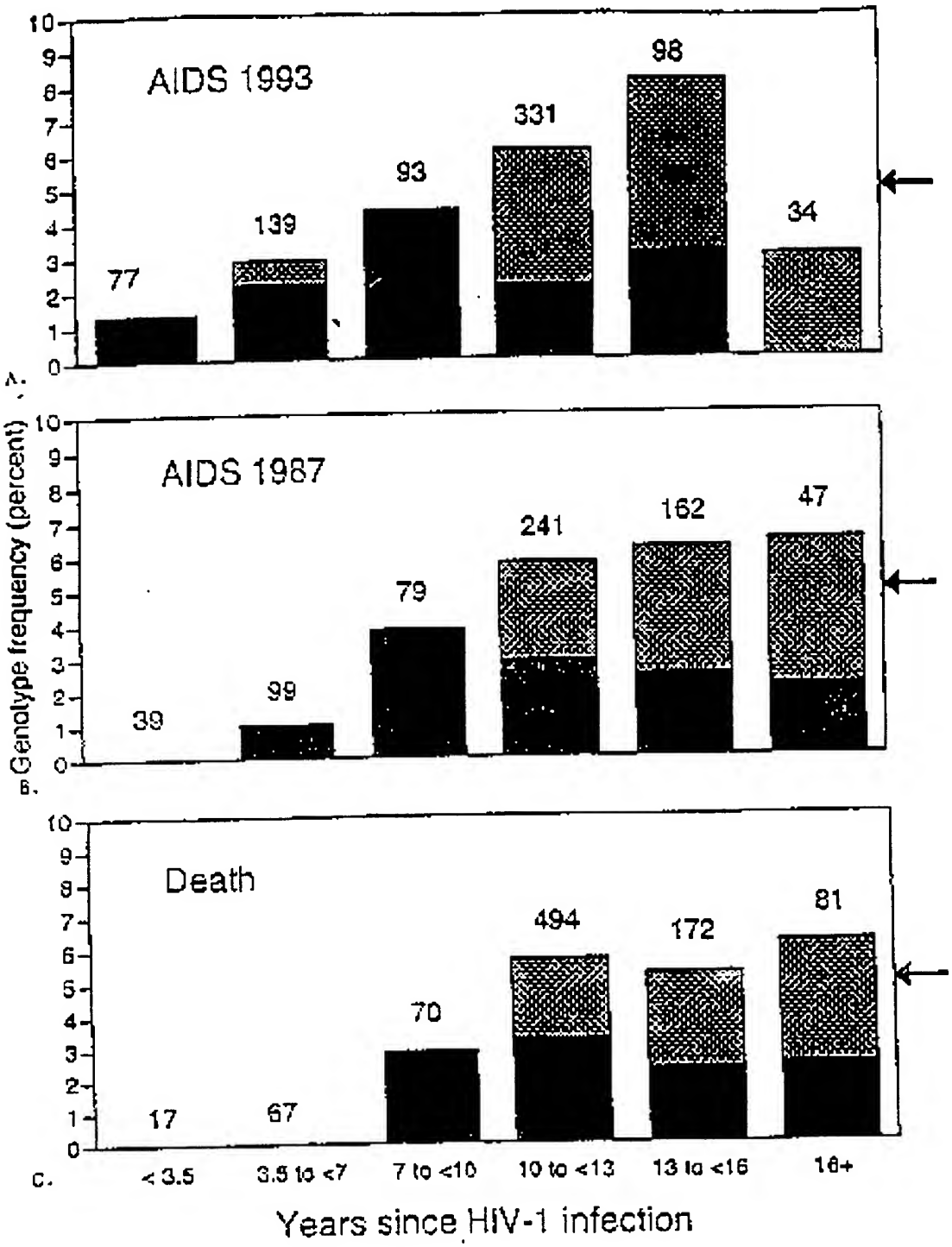
AIDS 1993
77
139
93
331
98
34
AIDS 1987
39
99
79
241
162
47
Death
17
67
70
494
172
81
Genotype frequency (percent)
A.
B.
C.
< 3.5
3.5 to <7
7 to <10
10 to <13
13 to <16
16+
Years since HIV-1 infection

FIGURE 5 A

SDF-1 3 - 3'A Mutation: (G801A) (SEQ ID NO:1)

```
   1 tctccgtcag ccgcattgcc cgctcggcgt ccggcccccg acccgtgctc gtccgccccc
  61 ccgccccgcc gcccgcgcca tgaacgccaa ggtcgtggtc gtgctggtcc tcgtgctgac
 121 cgcgctctgc ctcagcgacg ggaagcccgt cagcctgagc tacagatgcc catgccgatt
 181 cttcgaaagc catgttgcca gagccaacgt caagcatctc aaaattctca acactccaaa
 241 ctgtgccctt cagattgtag cccggctgaa gaacaacaac agacaagtgt gcattgaccc
 301 gaagctaaag tggattcagg agtacctgga gaaagctttg aacaagaggt tcaagatgtg
 361 agagggtcag acgcctgagg aaccccttaca gtaggagccc agctctgaaa ccagtgttag
 421 ggaagggcct gccacagcct cccctgccag ggcagggccc caggcattgc caagggcttt
 481 gttttgcaca ctttgccata ttttcaccat ttgattatgt agcaaaatac atgacattta
 541 ttttttcattt agtttgatta ttcagtgtca ctggcgacac gtagcagctt agactaaggc
 601 cattattgta cttgccttat tagagtgtct ttccacggag ccacccctct gactcagggc
 661 tcctgggttt tgtattctct gagctgtgca ggtgggggaga ctgggctgag ggagcctggc
 721 cccatggtca gcctagggt ggagagccac caagagggac gcctgggggt gccaggacca
 781 gtcaaccctgg gcaaagccta gtgaaggctt ctctctgtgg gatgggatgg tggagggcca
 841 catgggaggc tcaccccctt ctccatccac atgggagccg ggtctgcctc ttctgggagg
 901 gcagcagggc tacctgagc tgaggcagca gtgtgaggcc agggcagagt gagacccagc
 961 cctcatcccg agcacctcca catcctccac gttctgctca tcattctctg tctcatccat
1021 catcatgtgt gtccacgact gtctccatgg ccccgcaaaa ggactctcag gaccaaagct
1081 ttcatgtaaa ctgtgcacca agcaggaaat gaaaatgtct tgtgttacct gaaaacactg
1141 tgcacatctg tgtcttgttt ggaatattgt ccattgtcca atcctatgtt tttgttcaaa
1201 gccagcgtcc tcctctgtga ccaatgtctt gatgcatgca ctgtttcccc tgtgcagccg
1261 ctgagcgagg agatgctcct tgggcccttt gagtgcagtc ctgatcagag ccgtggtcct
1321 ttggggtgaa ctaccttggt tccccactg atcacaaaaa catggtgggt ccatgggcag
1381 agcccaaggg aattcgggtgt gcaccagggt tgaccccaga ggattgctgc cccatcagtg
1441 ctccctcaca tgtcagtacc tttaaactag ggccaagccc agcactgctt gaggaaaaca
1501 agcattcaca acttgttttt ggtttttaaa acccagtcca caaaataacc aatcctggac
1561 atgaagattc tttccaatt cacatctaac ctcatcttct tcaccatttg gcaatgccat
1621 catctcctgc cttcctcctg ggccctctct gctctgcgtg tcacctgtgc ttcgggccct
1681 tcccacagga catttctcta agaggacaat gtgctatgtg aagagtaagt caacctgccc
1741 gacatttgga gtgttccccc tccactgagg gcagtcgata gagctgtatt aagccactta
1801 aaatgttcac ttttgacaaa ggcaagcact tgtgggtttt tgttttgttt ttcattcagt
1861 cttcgaata cttttgcccc ttgattaaag actccagtta aaaaaaattt taatgaagaa
1921 agtggaaaac aaggaagtca aagcaaggaa actatgtaac atgtaggaag taggaagtaa
1981 attatagtga tgtaatcttg aattgtaact gttcttgaat ttaataatct gtagggtaat
2041 tagtaacatg tgttaagtat tttcataagt atttcaaatt ggagcttcat ggcagaaggc
2101 aaacccatca acaaaaattg tcccttaaac aaaaatttaaa atcctcaatc cagctatgtt
2161 atattgaaaa aatagagcct gagggatctt tactagttat aaagatacag aactctttca
2221 aaaccttttg aaattaacct ctcactatac cagtataatt gagttttcag tggggcagtc
2281 attatccagg taatccaaga tatttaaaa tctgtcacgt agaacttgga tgtacctgcc
2341 cccaatccat gaaccaagac cattgaattc ttggttgagg aaacaaacat gacccttaaat
2401 cttgactaca gtcaggaaag gaatcatttc tatttctcct ccatggagga aaatagataa
2461 gagtagaaac tgcagggaaa attatttgca taacaattcc tctactaaca atcagctcct
2521 cctggagaac tgcccagcta aagcaatatg catttaaata cagtctttca tttgcaaggg
2581 aaaagtctct tgtaatccga atctcttttt gctttcgaac tgctagtcaa gtgcgtccac
2641 gagctgttta ctagggatcc ctcatctgtc cctccgggac ctggtgctgc ctctaccttga
2701 cactcccttg ggctccctgt aacctcttca gaggccctcg ctgccagctc tgtatcaggg
2761 cccagaggaa ggggccagag gctcgttgac tggctgtgtg ttgggattga gtctgtgccg
2821 cgtgttttgtg ctgtggtgtg tccccctctg tccaggcact gagataccag cgaggaggct
2881 ccagagggca ctctgcttgt tattagagat tacctcctga gaaaaaagct tccgcttgga
2941 gcagaggggc tgaatagcag aaggttgcac ctccccaaac cttagatgtt ctaagtcttt
```